

# A Data-Scalable Randomized Misfit Approach for Solving Large-Scale PDE-Constrained Inverse Problems

E B Le<sup>1</sup>, A Myers<sup>1</sup>, T Bui-Thanh<sup>1,2</sup> and Q P Nguyen<sup>3</sup>

<sup>1</sup> Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, USA

<sup>2</sup> Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX, USA

<sup>3</sup> Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX, USA

E-mail: {ellenle,aaron,tanbui}@ices.utexas.edu and quoc\_p\_nguyen@mail.utexas.edu

**Abstract.** A randomized misfit approach is presented for the efficient solution of large-scale PDE-constrained inverse problems with high-dimensional data. The purpose of this paper is to offer a theory-based framework for random projections in this inverse problem setting. The stochastic approximation to the misfit is analyzed using random projection theory. By expanding beyond mean estimator convergence, a practical characterization of randomized misfit convergence can be achieved. The theoretical results developed hold with any valid random projection in the literature. The class of feasible distributions is broad yet simple to characterize compared to previous stochastic misfit methods. This class includes very sparse random projections which provide additional computational benefit. A different proof for a variant of the Johnson-Lindenstrauss lemma is also provided. This leads to a different intuition for the  $\mathcal{O}(\epsilon^{-2})$  factor in bounds for Johnson-Lindenstrauss results. The main contribution of this paper is a theoretical result showing the method guarantees a valid solution for small reduced misfit dimensions. The interplay between Johnson-Lindenstrauss theory and Morozov's discrepancy principle is shown to be essential to the result. The computational cost savings for large-scale PDE-constrained problems with high-dimensional data is discussed. Numerical verification of the developed theory is presented for model problems of estimating a distributed parameter in an elliptic partial differential equation. Results with different random projections are presented to demonstrate the viability and accuracy of the proposed approach.

*Keywords* Random projections, stochastic programming, large-scale inverse problems, model reduction, large deviation theory, Johnson-Lindenstrauss lemma, big data.

AMS classification scheme numbers: 35Q62, 62F15, 35R30, 35Q93, 65C60

## 1. Introduction

An emerging grand challenge in computational science and engineering is the solution of large-scale statistical inverse problems governed by PDEs that involve large amounts of observational data. These are difficult problems that can be found in diverse areas of science, engineering, and medicine, ranging from inference of the basal friction field in continental ice-sheet modeling to estimation of contaminant plume concentration in groundwater models. They are often characterized by infinite-dimensional parameter fields (for example, a spatially-distributed quantity) which results in a parameter dimension in the thousands or millions when discretized. We seek the Bayesian solution, as it offers a range of estimates that are consistent with data while accounting for uncertainty in the data, model, and prior knowledge. Unfortunately, this amounts to exploring a posterior probability, a task that is notoriously intractable for the problems of interest.

The dominant cost in this setting is measured in number of PDE solves. Each PDE solve takes minutes or hours even on modern supercomputers (e.g. [1–3]). State-of-the-art methods require repeated evaluations of an objective functions and its derivative information, resulting in a total cost of hundreds, thousands, or millions of PDE solves for many realistic problems. Reducing the number of PDE solves is paramount. Relative to the cost of a model run, linear algebra is considered negligible.

In this paper, we present a randomized misfit approach to directly address the computational burden induced by high-dimensional data. Note that the idea of randomizing a misfit function is not new. Randomized approximations of misfit functions can be found in methods for seismic inversion [4–6], in stochastic optimization algorithms such as stochastic gradient descent (see e.g. [7, 8]), and in the sample average approach (SAA) [9–11].

What is novel here is the particular randomized misfit framework and the resulting analysis. It is clear that the randomized objective function converges; it is less obvious that the minimizer converges. A connection with random projection theory is the key to understanding why the method results in an acceptable solution for a surprisingly small randomized misfit dimension, not just in the limit. This analysis potentially could be applied to existing methods that use randomized objective functions.

Roughly speaking, random projections are “quasi-orthogonal” transformations from high-dimensional spaces to much lower-dimensional spaces that, with high probability, preserve geometric properties such as Euclidean norms, distances, and angles. They are particularly revered for possessing such properties *independent of the original data dimension*. The geometric invariance properties are a consequence of the concentration of measure phenomenon in high dimensions. One can check that two high-dimensional random normal vectors on the unit sphere are nearly orthonormal, and that this phenomenon becomes more pronounced as the dimension grows larger. We show that for a broad class of distributions, the probability that a sample average falls within a specified ball around its mean grows exponentially high with the sample size. This

is the power of many independent random projections working together. Random projections provide probabilistic accuracy bounds that are parameterized by the degree of approximation or the dimension of the reduced space. That is, given a tolerance of approximation, one can find the reduced space dimension that will preserve Euclidean norm and vice versa. To assist in the practical use and verification of the method, in our numerical examples we use random projections that are easy to implement.

An active area of research is developing optimization methods for when the data set does not even fit in memory. The data needs to be subsampled prior to input. We stress here that this is not the main target of the randomized misfit approach. In the approach here, the data vector is not subsampled, but rather the misfit between the model and the data is linearly transformed to a smaller dimension where its geometric properties are preserved. This is equivalent to summing random linear combinations of the misfit components. We are not cleaning the data, fusing data points, or choosing a random subset of data to represent the full data set. We use the entire data set. The motivation is that the dominant cost in our problem setting is the number of PDE solves. The misfit vector dimension, as we will show, is a hard upper bound on a factor of the dominant cost. Thus if we can transform the misfit to a smaller dimension, we can reduce the dominant cost of solving the inverse problem and guarantee the accuracy of the solution. Computational cost is discussed in detail in Section 3.3.

The presentation is here is purposefully general and does not assume any particular underlying structure of the observational data, aside from its relationship to parameter space via the parameter-to-observable map and the noise model. Again there is a large body of work in data sampling, compression and/or fusion that exploits known underlying structure of the observational data set, typically for specific inverse problems. These methods are not incompatible with the approach we outline. They could potentially be combined with the method here to provide maximum computational savings.

### *1.1. Current state-of-the-art and our contributions*

To keep the discussion succinct and relevant to the problem of interest, where the dominant cost is PDE solves, the review is limited to existing work in randomized methods for PDE-constrained inverse problems. An active area of research is in applying random projections to linear regression problems. The dominant cost in these problems is generally measured in linear algebraic operations.

Since [12], many randomized methods to reduce the computational complexity of large-scale PDE-constrained inverse problems have focused on use of the randomized SVD algorithm of [13]. This algorithm has been used to generate truncated SVD approximations of the parameter-to-observable operator [14–18], the regularization operator [19,20], or the prior-preconditioned Hessian of the objective function [2,21–24]. The algorithm uses a random projection matrix to produce a low-rank operator. To our knowledge, only Gaussian distributions are used. The randomized operator is

subsequently factored to generate an approximate SVD decomposition for the original operator  $\mathcal{A}$ . Theoretical results in [13] guarantee the spectral norm accuracy of this approximation is of order  $\sigma_{k+1}(\mathcal{A})$  with a very high user-defined probability. Here  $k$  is equal to the reduced dimension  $n$  plus a small number of oversampling vectors. Subsequently, results known about the accuracy of a deterministic inverse solution (e.g., Proposition 1 in [25], Theorem 1 in [17]) to a problem approximated with a randomized method are derived using this bound from [13]. The bounds assume knowledge of  $\sigma_{k+1}(\mathcal{A})$ .

*Random source encoding* or *simultaneous (random) source* methods have been shown to be effective for parameter estimation in PDE-constrained inverse problems with multiple right-hand sides (sources) and corresponding data sets [6, 26–35]. This problem framework characterizes many inverse problems, including electromagnetic imaging (e.g. [36, 37]), seismic waveform inversion (e.g. [34, 38–40]), the DC resistivity problem (e.g. [41, 41]), and electromagnetic impedance tomography (e.g. [42] or Sec. 6.3 in [43]). Simultaneous source methods take random linear combinations of  $s$  sources to produce  $\tilde{s}$  randomly combined sources, where  $\tilde{s} \ll s$ . The result is a randomized misfit function that requires just  $\tilde{s}$  PDE solves to evaluate instead of  $s$  PDE solves. The work in [44] shows that source encoding in its stochastic reformulation (and as a stochastic trace estimator method [45]) is equivalent to an application of the random projection defined in [46]. Simultaneous source methods point out that numerical solutions are surprisingly better than the theory predicts with a small number of sources  $\tilde{s}$  (e.g.  $\tilde{s} \sim \mathcal{O}(1)$ ) [4, 26–30, 34, 35, 44].

This paper extends the above work in several directions. The approach outlined here allows for a stochastic reformulation of all PDE-constrained inverse problems recast in a constrained least-squares formulation, not just multi-source problems. Our analysis of computational efficiency is necessarily different and depends on how the large data dimension affects the optimization. The cost savings for this method is subtler than the cost savings in simultaneous source methods.

The main contribution of this paper is the first theoretical result to guarantee that the deterministic solution obtained with the randomized cost is valid for a fixed small reduced dimension. This provides an explanation for the surprising quality of solutions when using a randomized misfit function with a small reduced dimension. The efficacy of the randomized misfit approach is a result of the interplay between Morozov’s discrepancy principle and random projection theory. As will be shown, the data error and ill-posedness inherent in inverse problems is what allows random projections to be successful.

The key is the use of large deviation theory to arrive at a practical characterization of objective function convergence. Large deviation techniques form the basis of random projection theory and other effective randomized dimension reduction algorithms (e.g. randomized SVD). This approach exploits the concentration of measure phenomenon in high dimensions. We do not depend on the slow convergence of estimators to their exact mean. Instead it is shown that for a certain class of distributions, the tail probability of

a sample average of misfit estimators decays exponentially, with a rate parameterized by the sample size. This statement is turned into a probabilistic bound on the randomized cost for a fixed sample size.

The class with sufficient large deviation decay rate turns out to be *subgaussian* random variables. This class contains many of the distributions used in the simultaneous source literature. Another novel aspect of the approach here is that it permits the use of any random projection in the literature. Many new random projections have appeared since the seminal work in [46]. Potentially useful random projections are those drawn from very sparse distributions, which we test in our examples.

Additionally, from the stochastic formulation of the misfit, a different version of the Johnson-Lindenstrauss embedding theorem [47–49] is shown. This leads to an insight into why the reduced misfit dimension  $n$  is  $\mathcal{O}(\varepsilon^{-2})$ , where  $\varepsilon$  is the relative error of the randomized cost function.

The structure of the paper is as follows. Section 3 presents the theoretical analysis for the randomized misfit approach by deriving the large deviation bounds on the objective function error for a broad class of distributions. The reduced misfit dimension is shown to be independent of the original data dimension. This derivation leads to a different proof of a variant of the celebrated Johnson-Lindenstrauss embedding theorem. Using Morozov’s discrepancy principle, Theorem 2 shows that the effective reduced misfit dimension is also bounded below by the noise in the problem. Therefore, the RMA solution is a guaranteed solution for the original problem with a high user-defined probability. The reduced computational cost in problems with high-dimensional data is assessed in Section 3.3. Section 4 summarizes numerical experiments on a model inverse heat conduction problem in one-, two-, and three spatial-dimensions. We compare the RMA solution obtained with different distributions to the solution of the full problem. We also provide numerical support for Theorem 2.

## 2. The randomized misfit approach for inverse problems

We assume an additive noise-corrupted pointwise observational model

$$d_j = w(\mathbf{x}_j; u) + \eta_j, \quad j = 1, \dots, N, \quad (1)$$

where the objective is to reconstruct the distributed parameter  $u$  given  $N$  data points  $d_j$ , with  $N$  large. For a given  $u$ , a set of states  $w(\mathbf{x}_j; u)$  is obtained by evaluating an expensive-to-solve forward model governed by PDEs, and then applying a linear observation operator to match the data locations. The location of an observational data point in an open and bounded spatial domain  $\Omega$  is denoted by  $\mathbf{x}_j$ , and  $\eta_j$  is assumed to be Gaussian random noise with mean 0 and variance  $\sigma^2$ .

Concatenating the observations, we rewrite (1) as

$$\mathbf{d} = \mathbf{F}(u) + \boldsymbol{\eta}, \quad (2)$$

where  $\mathbf{F}(u) := [w(\mathbf{x}_1; u), \dots, w(\mathbf{x}_N; u)]^\top$  is the parameter-to-observable map. Although the forward problem is usually well-posed, the inverse problem is ill-posed. An intuitive reason is that discrete observations can only contain limited information about an infinite-dimensional parameter. The more complete explanation is that the parameter-to-observable map exhibits rapid spectral decay. This can be numerically observed and is proven for many practical inverse problems [50–52]. By an application of Picard’s Theorem we may then show the inverse operator is unbounded and therefore the problem is ill-posed.

A standard deterministic Tikhonov approach resolves the ill-conditioning by adding a quadratic term to the cost function, so that the problem may now be formulated as

$$\min_u \mathcal{J}(u) := \frac{1}{2} \left\| \hat{\mathbf{d}} - \hat{\mathbf{F}}(u) \right\|^2 + \frac{1}{2} \left\| R^{\frac{1}{2}} u \right\|^2, \quad (3)$$

where  $\hat{\mathbf{d}} - \hat{\mathbf{F}}(u) := \frac{1}{\sigma} [\mathbf{d} - \mathbf{F}(u)]$  is the *data misfit vector*, Euclidean norm in  $\mathbb{R}^N$  is denoted by  $\|\cdot\|$ , and  $\left\| R^{\frac{1}{2}} \cdot \right\|$  is a norm weighted by a regularization operator  $R$ . This point estimate does not account for the uncertainty in the solution. Thus we recast the problem in the framework of Bayesian inference, where we seek a statistical description of all possible parameter fields that are consistent with the observations. The Bayesian solution is a probability distribution that accounts for the uncertainties in the observations, the forward model, and the prior knowledge. It requires specification of a likelihood model, which characterizes the probability that the parameter  $u$  could have produced the observed data  $\mathbf{d}$ . It also requires a prior model, which is problem-dependent and represents a subjective belief regarding the distribution of  $u$ . The prior model must ensure sufficient regularity of the parameter so that the problem is well-posed [21, 53, 54].

The additive-noise model (2) is used to construct the likelihood pdf which is expressed as

$$\pi_{\text{like}}(\mathbf{d}|u) \propto \exp \left( -\frac{1}{2} \left\| \hat{\mathbf{d}} - \hat{\mathbf{F}}(u) \right\|^2 \right). \quad (4)$$

For concreteness of presentation, we postulate that the prior is a Gaussian random field with mean  $u_0$  and a covariance operator  $\mathcal{C}$ . We must stress here that the choice of a meaningful prior in the infinite-dimensional setting is an active area of research [53, 55–57]. The Gaussian prior used here is chosen only to ensure well-posedness and be computationally amenable to general large-scale problems. We choose  $\mathcal{C} = \mathcal{A}^{-2}$ , where  $\mathcal{A}$  is a Laplacian-like operator with its domain of definition specified by an elliptic PDE, appropriately-chosen boundary conditions, and parameters that can encode spatial correlation and anisotropy information (for specific implementation details see [14, 21]). This choice avoids constructing and inverting a dense covariance matrix and exploits existing fast solvers for elliptic operators. It additionally provides a connection to the Matérn covariance functions used frequently in geostatistics [22, 58, 59] and therefore has a scientific justification. Note that directly specifying a covariance function and then factorizing the covariance matrix is common and reasonable for small- to medium-scale statistical inverse problems, but is intractable for large-scale problems [3].

We then discretize the prior, the forward equation, and the parameter  $u$  (yielding a finite-dimensional vector  $\mathbf{u} \in \mathbb{R}^m$ ) through the finite element method (see [21, 54] for a comprehensive treatment) so that the finite-dimensional posterior probability of  $\mathbf{u}$  is given by Bayes formula as

$$\pi(\mathbf{u}|\mathbf{d}) \propto \exp\left(\frac{1}{2}\left\|\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right\|^2 + \frac{1}{2}\|\mathbf{u} - \mathbf{u}_0\|_{\mathcal{C}}^2\right) \quad (5)$$

where  $\|\cdot\|_{\mathcal{C}} := \left\|\mathcal{C}^{-\frac{1}{2}}\cdot\right\|_{L^2(\Omega)}$  denotes the weighted  $L^2(\Omega)$  norm induced by the  $L^2(\Omega)$  inner product  $\langle \cdot, \cdot \rangle_{L^2(\Omega)}$ . The maximum a posteriori (MAP) point of (5) is defined as

$$\mathbf{u}^* := \arg \min_{\mathbf{u}} J(\mathbf{u}, \mathbf{d}) = \frac{1}{2}\left\|\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right\|^2 + \frac{1}{2}\|\mathbf{u} - \mathbf{u}_0\|_{\mathcal{C}}^2. \quad (6)$$

Note that the last term in (6) may be viewed as a Tikhonov regularization term, and subsequently the MAP point may be considered as a solution to the deterministic inverse problem with a regularization “inspired” by the prior. Understanding the MAP point in a Bayesian framework allows one to account for the subjectivity of choosing a prior. Ultimately, the goal is to find the Bayesian solution which offers a statistical description of *all solutions* consistent with the data. For this paper, we restrict ourselves to MAP computation, a necessary starting point, in order to focus on methodology development in addressing the challenge of big data, i.e., large  $N$ . Scalability and efficiency of the method in the Bayesian setting is the focus of ongoing work.

The main idea of the randomized misfit approach is the following. Let  $\mathbf{r} \in \mathbb{R}^N$  be a random vector with mean zero and identity covariance, i.e.  $\mathbb{E}_{\mathbf{r}}[\mathbf{r}\mathbf{r}^\top] = \mathbf{I}$  (equivalently, let  $\mathbf{r}$  be the vector of  $N$  i.i.d. random variables  $\zeta$  with mean zero and variance 1).

Then the misfit term of (6) can be rewritten as:

$$\left\|\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right\|^2 = \left(\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right)^\top \mathbb{E}_{\mathbf{r}}[\mathbf{r}\mathbf{r}^\top] \left(\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right) = \mathbb{E}_{\mathbf{r}}\left[\mathbf{r}^\top \left(\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right)\right]^2, \quad (7)$$

which allows us to write the objective functional in (6) as

$$J(\mathbf{u}) = \frac{1}{2}\mathbb{E}_{\mathbf{r}}\left[\mathbf{r}^\top \left(\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right)\right]^2 + \frac{1}{2}\|\mathbf{u} - \mathbf{u}_0\|_{\mathcal{C}}^2. \quad (8)$$

We then approximate the expectation  $\mathbb{E}_{\mathbf{r}}[\cdot]$  using a Monte Carlo approximation (also known as the Sample Average Approximation (SAA) [9, 10]) with  $n$  i.i.d. draws  $\{\mathbf{r}_j\}_{j=1}^n$ . This leads to the randomized inverse problem

$$\begin{aligned} \min_{\mathbf{u}} J_n(\mathbf{u}; \mathbf{r}) &= \frac{1}{2n} \sum_{j=1}^n \left[\mathbf{r}_j^\top \left(\widehat{\mathbf{d}} - \widehat{\mathbf{F}}(\mathbf{u})\right)\right]^2 + \frac{1}{2}\|\mathbf{u} - \mathbf{u}_0\|_{\mathcal{C}}^2 \\ &= \frac{1}{2}\left\|\tilde{\mathbf{d}} - \tilde{\mathbf{F}}(\mathbf{u})\right\|^2 + \frac{1}{2}\|\mathbf{u} - \mathbf{u}_0\|_{\mathcal{C}}^2, \end{aligned} \quad (9)$$

where  $\tilde{\mathbf{d}} := \frac{1}{\sqrt{n}}[\mathbf{r}_1, \dots, \mathbf{r}_n]^\top \widehat{\mathbf{d}}$ , and  $\tilde{\mathbf{F}}(\mathbf{u}) := \frac{1}{\sqrt{n}}[\mathbf{r}_1, \dots, \mathbf{r}_n]^\top \widehat{\mathbf{F}}(\mathbf{u}) \in \mathbb{R}^n$ . We call  $\tilde{\mathbf{d}} - \tilde{\mathbf{F}}(\mathbf{u})$  the *reduced data misfit vector*.

For a *reduced misfit vector dimension*  $n \ll N$ , we call this randomization the randomized misfit approach (RMA). The new problem (9) with fixed i.i.d. realizations  $\{\mathbf{r}_j\}_{j=1}^n$  may be solved using any scalable robust optimization algorithm. For the numerical experiments in Section 4, a globalized inexact Newton-CG implementation [60] is used. The use of a similar mesh-independent Newton-type method is assumed for the discussion of computational complexity in Section 3.3.

We define the MAP point of (9) as

$$\mathbf{u}_n^* := \arg \min_{\mathbf{u}} J_n(\mathbf{u}), \quad (10)$$

the optimal RMA cost as  $J_n^* := J_n(\mathbf{u}_n^*)$ , and the optimal true cost as  $J^* := J(\mathbf{u}^*)$ . We wish to characterize the errors  $|J_n^* - J^*|$  and  $\|\mathbf{u}_n^* - \mathbf{u}^*\|$  for a given reduced misfit dimension  $n$ . This is the subject of section 3.

### 3. An analysis of the randomized misfit approach (RMA)

#### 3.1. Validity of the RMA solution

For a given  $\mathbf{u}$  in parameter space, it is clear that  $J_n(\mathbf{u}; \mathbf{r})$  in (9) is an unbiased estimator of  $J(\mathbf{u})$ . It is also clear from the Law of Large Numbers that  $J_n(\mathbf{u})$  converges almost surely to its mean  $J(\mathbf{u})$ . However, the efficacy of the randomized misfit approach lies in exploiting the concentration of measure phenomenon of high dimensions, and quantifying the convergence *close to the mean*. This requires characterizing the exponential decay of the objective function error, which is parameterized by the reduced misfit dimension  $n$ .

We first show that errors larger than  $\delta/2$ , for a given  $\delta > 0$ , decay with a rate at least as fast as the tail of a centered Gaussian. That is, for some distribution in (9) we have

$$\mathbb{P} \left[ |J_n(\mathbf{u}; \mathbf{r}) - J(\mathbf{u})| > \frac{\delta}{2} \right] \leq e^{-nI(\delta)}, \quad (11)$$

where

$$I(\delta) \geq c \frac{\delta^2}{2\theta^2}. \quad (12)$$

for some  $c > 0$  and some  $\theta$ .

This rate is sufficient to guarantee the solution attained from the the randomized misfit approach is a discrepancy principle-satisfying solution for the original inverse problem as will be shown in Theorem 2. Inequality (11) is equivalent to the statement that  $\mathbb{P} [|J_n(\mathbf{u}; \mathbf{r}) - J(\mathbf{u})| > \frac{\delta}{2}]$  satisfies a *large deviation principle* with *large deviation rate function*  $I(\delta)$  [61].

The following proposition may be viewed as a special case of Cramér's Theorem, which states that a sample mean of i.i.d. random variables  $X$  asymptotically obeys a large deviation principle with rate  $I(\delta) = \sup_k \{k\delta - \ln \mathbb{E}[e^{kX}]\}$  [61]. However we require the exact non-asymptotic bounds as derived here to show convergence of the



RMA for  $n = \mathcal{O}(1)$ . Recall that a real-valued random variable  $X$  is  $\theta$ -subgaussian if there exists some  $\theta > 0$  such that for all  $t \in \mathbb{R}$ ,  $\mathbb{E}[e^{tX}] \leq e^{\theta^2 t^2/2}$ .

**Proposition 1** *The RMA error  $|J_n(\mathbf{u}; \mathbf{r}) - J(\mathbf{u})|$  has a tail probability that decays exponentially in  $n$  with a nontrivial large deviation rate. Furthermore, if the RMA is constructed with  $\mathbf{r}$  such that  $2|J_n(\mathbf{u}; \mathbf{r}) - J(\mathbf{u})|$  is the sample mean of i.i.d.  $\theta$ -subgaussian random variables, then its large deviation rate is bounded below by  $c \frac{\delta^2}{2\theta^2}$  for some  $c > 0$ .*

*Proof.* Given  $\mathbf{r}$ , define the random variable

$$T(\mathbf{r}; \mathbf{u}) := \left[ \mathbf{r}^\top (\hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u})) \right]^2 - \left\| \hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u}) \right\|^2. \quad (13)$$

By a standard Chernoff bound (see, e.g. [62]), we have that the RMA tail error decays exponentially as

$$\mathbb{P} \left[ \frac{1}{n} \sum_{j=1}^n T(\mathbf{r}_j; \mathbf{u}) > \delta \right] \leq e^{-nI(\delta)}, \quad (14)$$

where  $I(\delta) = \max_t \{t\delta - \ln \mathbb{E}[e^{tT(\mathbf{r}; \mathbf{u})}]\}$  is the large deviation rate.

The second part of the proposition follows with  $c = 1$  by bounding  $\mathbb{E}[e^{tT(\mathbf{r}; \mathbf{u})}]$  in (14) and computing the maximum of  $t\delta - \theta^2 t^2/2$ .  $\square$

A large number of distributions are subgaussian, notably the Gaussian and Rademacher (also referred to as Bernoulli) distributions, and in fact any bounded random variable is subgaussian. One class of subgaussian distributions that provides additional computational efficiency is the following.

**Definition 1** ( $\ell$ -percent sparse random variables [48, 63]) *Let  $s = \frac{1}{1-\ell}$  where  $\ell \in [0, 1)$  is the level of sparsity desired. Then*

$$\zeta = \sqrt{s} \begin{cases} +1 & \text{with probability } \frac{1}{2s}, \\ 0 & \text{with probability } \ell = 1 - \frac{1}{s}, \\ -1 & \text{with probability } \frac{1}{2s} \end{cases} \quad (15)$$

*is a  $\ell$ -percent sparse distribution.*

Note that for  $\ell = 0$ ,  $\zeta$  corresponds to a Rademacher distribution, and that  $\ell = 2/3$  corresponds to the *Achlioptas distribution* [46]. By inspection we have that  $\mathbb{E}[\zeta] = 0$  and  $\mathbb{E}[\zeta^2] = 1$ , and thus draws from  $\zeta$  can be used in the randomized misfit approach.

Distribution (15) is well-suited for the randomized misfit approach: it is easy to implement, and the computation of the randomized misfit vector amounts to only summations and subtractions, adding a further speedup to the method. Increasing from  $s = 1$  to  $s > 1$  results in a  $s$ -fold speedup as only  $1/s$  of the data is included. Note the RMA cost can be seen as the sum of  $n$  random combinations from the  $N$ -dimensional misfit vector. Since each random combination has a different sparsity pattern, we effectively do not exclude any data, yet each computation requires only  $1/s$  of the data.

We note that for the distribution (15),  $1 \leq s < \infty$ , the random variable  $\zeta$  distributed by (15) has<sup>†</sup>  $\mathbb{E}[e^{t\zeta}] \leq e^{\frac{b^2 t^2}{2}}$  with  $b = \sqrt{s - 2 \ln s}$ ,  $\forall t \in (0, 1]$ . So, we may use it in the following theorem.

**Theorem 1** Define  $\mathbf{v} := \hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u}) \in \mathbb{R}^N$ . If  $\mathbf{r}$  in (13) has components that are  $b$ -subgaussian for some  $b \geq 1/\sqrt{2}$ , then the RMA error has a large deviation rate bounded below by  $c \frac{\delta^2}{2\theta^2}$  for  $\theta = \|\mathbf{v}\|^2 / \sqrt{2}$  and some  $0 < c < \frac{1}{8b^4}$ .

*Proof.* Let  $\mathbf{r} \in \mathbb{R}^N$  such that  $\mathbf{r}$  has i.i.d.  $b$ -subgaussian components  $r_i$ , with  $b \geq 1/\sqrt{2}$ ,  $\mathbb{E}[r_i] = 0$ , and  $\mathbb{E}[r_i^2] = 1$ . Define  $\mathbf{w} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$  and  $X = \mathbf{r}^\top \mathbf{w}$ . Then

$$\mathbb{E}[e^{tT}] = e^{-t\|\mathbf{v}\|^2} \mathbb{E}[e^{t\|\mathbf{v}\|^2 X^2}] \quad \forall t \in \mathbb{R}. \quad (17)$$

From [48, Lemma 2.2],  $\mathbb{E}[X^2] = 1$  and  $X$  is also  $b$ -subgaussian. Then, by [49, Remark 5.1], for  $0 \leq t \leq \frac{1}{4b^2}$ ,

$$\mathbb{E}[e^{tX^2}] \leq \sqrt{2}. \quad (18)$$

For  $0 < t \leq \frac{1}{4b^2\|\mathbf{v}\|^2}$ , we have

$$\begin{aligned} \mathbb{E}[e^{t\|\mathbf{v}\|^2 X^2}] &\leq 1 + t\|\mathbf{v}\|^2 + t\|\mathbf{v}\|^4 \frac{\mathbb{E}[X^4]}{2} + \sum_{k=3}^{\infty} \frac{\left(\frac{1}{4b^2}\right)^k (4b^2 t \|\mathbf{v}\|^2)^k \mathbb{E}[X^{2k}]}{k!} \\ &\leq 1 + t\|\mathbf{v}\|^2 + t^2\|\mathbf{v}\|^4 \frac{\mathbb{E}[X^4]}{2} + (4b^2 t \|\mathbf{v}\|^2)^3 \sum_{k=3}^{\infty} \frac{\left(\frac{1}{4b^2}\right)^k \mathbb{E}[X^{2k}]}{k!} \\ &\leq 1 + t\|\mathbf{v}\|^2 + t^2\|\mathbf{v}\|^4 \frac{\mathbb{E}[X^4]}{2} + (4b^2 t \|\mathbf{v}\|^2)^3 \mathbb{E}\left[e^{\frac{1}{4b^2} X^2}\right] \\ &\leq 1 + t\|\mathbf{v}\|^2 + t^2\|\mathbf{v}\|^4 \frac{\mathbb{E}[X^4]}{2} + 64\sqrt{2}b^6 t^3 \|\mathbf{v}\|^6 \\ &\leq 1 + t\|\mathbf{v}\|^2 + 8b^4 t^2 \|\mathbf{v}\|^4 + 64\sqrt{2}b^6 t^3 \|\mathbf{v}\|^6 \\ &\leq e^{t\|\mathbf{v}\|^2 + 8b^4 t^2 \|\mathbf{v}\|^4 + 64\sqrt{2}b^6 t^3 \|\mathbf{v}\|^6}, \end{aligned}$$

using (18) in the fourth inequality and [64, p.93] in the fifth inequality. Let  $t_\star = \frac{\delta}{8b^4\|\mathbf{v}\|^4 q}$  where  $q > 1$ . Assuming  $\|\mathbf{v}\|^2 \gg \delta$ , we have that

$$\mathbb{E}[e^{t_\star T}] \leq e^{8b^4 t_\star^2 \|\mathbf{v}\|^4 + 64\sqrt{2}b^6 t_\star^3 \|\mathbf{v}\|^6} = e^{\frac{\delta^2}{8b^4\|\mathbf{v}\|^4 q^2} + \sqrt{2}\frac{\delta^3}{8b^6\|\mathbf{v}\|^6 q^3}}.$$

Then

$$I(\delta) \geq \delta t_\star - \ln \mathbb{E}[e^{t_\star T}] \geq \left(1 - \frac{1}{q}\right) \frac{\delta^2}{8b^4\|\mathbf{v}\|^4 q} - \sqrt{2} \frac{\delta^3}{8b^6\|\mathbf{v}\|^6 q^3} \geq c \frac{\delta}{\|\mathbf{v}\|^4},$$

<sup>†</sup> Using the inequality  $(2k)! \geq 2^k k!$  and the Taylor expansion around 0, we have that for  $t \in (0, 1]$

$$\mathbb{E}[e^{t\zeta}] = \frac{1}{s} \sum_{k=0}^{\infty} \frac{(st^2)^k}{(2k)!} \leq \frac{1}{s} \sum_{k=0}^{\infty} \frac{(st^2)^k}{2^k k!} = \frac{1}{s} e^{\frac{s}{2} t^2} = e^{-\ln s + \frac{s}{2} t^2} \leq e^{-t^2 \ln s + \frac{s}{2} t^2}. \quad (16)$$

where  $0 < c < \frac{1}{8b^4}$ . Taking  $2\theta^2 = \|\mathbf{v}\|^4$  concludes the proof.  $\square$

A sharper result can be obtained for RMA constructed with  $b$ -subgaussian random variables where  $b \leq 1$ . Note that this includes the distribution (15) with  $s = 1$  (Rademacher) and  $s = 3$  (Achlioptas) by the above theorem. Following [49, (5)], let  $g$  be a standard Gaussian random variable, independent of all other random variables. Then, we have that for  $0 < t < \frac{1}{2\|\mathbf{v}\|^2}$ ,

$$\mathbb{E} \left[ e^{t\|\mathbf{v}\|^2 X} \right] \leq \mathbb{E}_g \left[ \prod_i^N e^{b^2 t \|\mathbf{v}\|^2 w_i^2 g^2} \right] \leq \mathbb{E}_g \left[ e^{t\|\mathbf{v}\|^2 g^2} \right] = \frac{1}{\sqrt{1 - 2t\|\mathbf{v}\|^2}}. \quad (19)$$

So from (17) we have that

$$\mathbb{E} \left[ e^{tT(\mathbf{u}, \mathbf{r})} \right] \leq \frac{e^{-t\|\mathbf{v}\|^2}}{\sqrt{1 - 2t\|\mathbf{v}\|^2}} = e^{-t\|\mathbf{v}\|^2 - \frac{1}{2} \ln(1 - 2t\|\mathbf{v}\|^2)}. \quad (20)$$

Then

$$t\delta - \ln(\mathbb{E}[T(\mathbf{u}, \mathbf{r})]) \geq t\delta + t\|\mathbf{v}\|^2 + \frac{1}{2} \ln(1 - 2t\|\mathbf{v}\|^2) =: f(t). \quad (21)$$

Computing the derivative, we have that  $f(t)$  attains a maximum at

$$t_{\max} = \frac{\delta}{2(\|\mathbf{v}\|^4 + \delta\|\mathbf{v}\|^2)}. \quad (22)$$

Thus, we have

$$\begin{aligned} \max f(t) &= \frac{\delta^2}{2(\|\mathbf{v}\|^4 + \delta\|\mathbf{v}\|^2)} + \frac{\delta}{2(\|\mathbf{v}\|^2 + \delta)} + \frac{1}{2} \ln \left( 1 - \frac{\delta}{\|\mathbf{v}\|^2 + \delta} \right) \\ &= \frac{\delta^2}{2(\|\mathbf{v}\|^4 + \delta\|\mathbf{v}\|^2)} - \frac{1}{4} \frac{\delta^2}{(\|\mathbf{v}\|^2 + \delta)^2} - \frac{1}{6} \frac{\delta^3}{(\|\mathbf{v}\|^2 + \delta)^3} - \dots \\ &= \frac{\delta^2}{4(\|\mathbf{v}\|^4 + \delta\|\mathbf{v}\|^2)} + \frac{1}{4} \left\{ \frac{\delta^2}{(\|\mathbf{v}\|^4 + \delta\|\mathbf{v}\|^2)} - \frac{\delta^2}{(\|\mathbf{v}\|^2 + \delta)^2} \right\} \\ &\quad - \frac{1}{6} \frac{\delta^3}{(\|\mathbf{v}\|^2 + \delta)^3} - \dots \geq c \frac{\delta^2}{\|\mathbf{v}\|^4}, \end{aligned}$$

where we employed the Taylor expansion in the second equality, and in the last inequality  $c$  is some constant less than  $1/4$ . Note that the last inequality holds for  $\delta \ll \|\mathbf{v}\|^2$  and taking  $2\theta^2 = \|\mathbf{v}\|^4$  concludes the proof.

The next theorem is our main result. It guarantees with high probability that the RMA solution will be a solution of the original problem under Morozov's discrepancy principle, for relatively small  $n$ . We first need the following lemma.

**Lemma 1** *Let  $\mathbf{v} := \hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u})$ . Suppose that  $\mathbf{r}$  is distributed such that the large deviation rate of the RMA error is bounded below by  $c \frac{\delta^2}{2\theta^2}$  for some  $c > 0$  and  $\theta = \|\mathbf{v}\|^2 / \sqrt{2}$ . Given a cost distortion tolerance  $\varepsilon > 0$  and a failure rate  $\beta > 0$ , let*

$$n \geq \frac{\beta}{c\varepsilon^2}. \quad (23)$$

Then with probability at least  $1 - e^{-\beta}$ ,

$$(1 - \varepsilon) \|\mathbf{v}\|^2 \leq \frac{1}{n} \sum_{j=1}^n (\mathbf{r}_j^\top \mathbf{v})^2 \leq (1 + \varepsilon) \|\mathbf{v}\|^2, \quad (24)$$

and hence,

$$(1 - \varepsilon) J(\mathbf{u}) \leq J_n(\mathbf{u}; \mathbf{r}) \leq (1 + \varepsilon) J(\mathbf{u}). \quad (25)$$

*Proof.* The proof follows from setting  $\delta = \varepsilon \|\mathbf{v}\|^2$  in (11).  $\square$

This lemma demonstrates a remarkable fact that with  $n$  i.i.d. draws one can reduce the data misfit dimension from  $N$  to  $n$  while bearing a relative error of  $\varepsilon = \mathcal{O}(1/\sqrt{n})$  in the cost function, where the reduced dimension  $n$  is independent of the dimension  $N$  of the data. This idea is the basis for data-reduction techniques via variants of the Johnson-Lindenstrauss Lemma in existing work with random projections (see e.g. [65–67]). With the connection through the randomized misfit approach, the ubiquitous  $N$ -independent Monte Carlo factor  $\varepsilon = \mathcal{O}(1/\sqrt{n})$  in Johnson-Lindenstrauss literature can thus be understood by reframing the application of a random projection as a Monte Carlo method in the form of (24).

Unlike other applications of the Monte Carlo method, e.g. Markov chain Monte Carlo, in which  $n$  must be large to be successful,  $n$  can be moderate or small for inverse problems, depending on the noise  $\boldsymbol{\eta}$  in (2). In the following theorem we show this is possible via Morozov's discrepancy principle [68]. To avoid over-fitting the noise, from (1) one seeks a MAP point  $\mathbf{u}^*$  such that  $|d_j - w(\mathbf{x}_j; \mathbf{u}^*)| \approx \sigma$ , i.e.  $\|\hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u}^*)\|^2 \approx N$ . We say that an inverse solution  $\mathbf{u}^*$  satisfies Morozov's discrepancy principle with parameter  $\tau$  if

$$\|\hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u}^*)\|^2 = \tau N \quad (26)$$

for some  $\tau \approx 1$ .

**Theorem 2 (Statistical Morozov's discrepancy principle)** *Suppose that the conditions of Lemma 1 are met. If  $\mathbf{u}_n^*$  is a discrepancy principle-satisfying solution for the RMA cost, i.e.,*

$$\mathcal{J}_n(\mathbf{u}_n^*, \mathbf{r}) := \frac{1}{n} \sum_{j=1}^n \left[ \mathbf{r}_j^\top (\hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u}_n^*)) \right]^2 = \tau' N \quad (27)$$

*for some  $\tau' \approx 1$ , then with probability at least  $1 - e^{-\beta}$ ,  $\mathbf{u}_n^*$  is also a solution for the original problem that satisfies Morozov's discrepancy principle with parameter  $\tau$ , i.e.*

$$\mathcal{J}(\mathbf{u}_n^*) := \|\hat{\mathbf{d}} - \hat{\mathbf{F}}(\mathbf{u}_n^*)\|^2 = \tau N. \quad (28)$$

*for  $\tau \in [\frac{\tau'}{1+\varepsilon}, \frac{\tau'}{1-\varepsilon}]$ .*

*Proof.* The claim is a direct consequence of (24).  $\square$

### 3.2. Other theoretical results

We are now in the position to show a different proof of the Johnson-Lindenstrauss embedding theorem using a stochastic programming derivation of the RMA. Following [69], we define a map  $\mathcal{S}$  from  $\mathbb{R}^n$  to  $\mathbb{R}^N$ , where  $n \ll N$ , to be a Johnson-Lindenstrauss transform (JLT) if

$$(1 - \varepsilon) \|\mathbf{v}\|^2 \leq \|\mathcal{S}\mathbf{v}\|^2 \leq (1 + \varepsilon) \|\mathbf{v}\|^2, \quad (29)$$

holds with some probability  $p = p(n, \varepsilon)$ , where  $\varepsilon > 0$ .

**Theorem 3 (Johnson-Lindenstrauss embedding theorem [47–49])** *Suppose that  $\mathbf{r}$  is distributed such that the large deviation rate of the RMA error is bounded below by  $c \frac{\delta^2}{2\theta^2}$  for some  $c > 0$  and some  $\theta$ . Let  $0 < \varepsilon < 1$ ,  $\mathbf{v}_i \in \mathbb{R}^N, i = 1, \dots, m$ , and  $n = \mathcal{O}(\varepsilon^{-2} \ln m)$ . Then there exists a map  $\mathcal{F} : \mathbb{R}^N \rightarrow \mathbb{R}^n$  such that*

$$(1 - \varepsilon) \|\mathbf{v}_i - \mathbf{v}_j\|^2 \leq \|\mathcal{F}(\mathbf{v}_i) - \mathcal{F}(\mathbf{v}_j)\|^2 \leq (1 + \varepsilon) \|\mathbf{v}_i - \mathbf{v}_j\|^2 \quad \forall i, j. \quad (30)$$

*Proof.* The conditions of Lemma 1 hold, thus for a given  $\mathbf{v} \in \mathbb{R}^N$ , note that (24) is equivalent to

$$(1 - \varepsilon) \|\mathbf{v}\|^2 \leq \|\Sigma \mathbf{v}\|^2 \leq (1 + \varepsilon) \|\mathbf{v}\|^2, \quad (31)$$

where

$$\Sigma := \frac{1}{\sqrt{n}} [\mathbf{r}_1, \dots, \mathbf{r}_n]^\top. \quad (32)$$

Define  $\mathcal{F}(\mathbf{v}) := \Sigma \mathbf{v}$ . Inequality (30) is then a direct consequence of (31) for a pair  $(\mathbf{v}_i, \mathbf{v}_j)$  with probability at least  $1 - e^{-\frac{\varepsilon}{2} n \varepsilon^2}$ . Using an union bound over all pairs, claim (30) holds for any pair with probability at least  $1 - m^{-\alpha}$  if  $n \geq c \frac{(2+\alpha)}{\varepsilon^2} \ln m$ .  $\square$

As discussed above,  $J_n(\mathbf{u}; \mathbf{r})$  is an unbiased estimator of  $J(\mathbf{u})$ . It is therefore reasonable to expect that  $J_n^* := \min_{\mathbf{u}} J_n(\mathbf{u}; \mathbf{r})$  converges to  $J^* := \min_{\mathbf{u}} J(\mathbf{u})$ . The following result [10, Propositions 5.2 and 5.6] states that under mild conditions  $J_n^*$  in fact converges to  $J^*$ . It is not unbiased, but is however downward biased.

**Proposition 2** *Assume that  $J_n(\mathbf{u}; \mathbf{r})$  converges to  $J(\mathbf{u})$  with probability 1 uniformly in  $\mathbf{u}$ , then  $J_n^*$  converges to  $J^*$  with probability 1. Furthermore, it holds that*

$$\mathbb{E}[J_n^*] \leq \mathbb{E}[J_{n+1}^*] \leq J^*, \quad (33)$$

that is,  $J_n^*$  is a downward-biased estimator of  $J^*$ .

Stochastic programming theory gives a stronger characterization of this convergence. One can show that  $\mathbf{u}_n^*$  converges weakly to  $\mathbf{u}^*$  with an  $n^{-\frac{1}{2}}$  rate. If  $J(\mathbf{u})$  is convex with finite value, then  $\mathbf{u}_n^* = \mathbf{u}^*$  with probability exponentially converging to 1. See Chapter 5 in [10] for details. For a linear forward map  $\mathbf{F}(\mathbf{u}) = \mathbf{F}\mathbf{u}$ , that is,  $J(\mathbf{u})$  is quadratic, we can derive a bound on the solution error using the spectral norm of  $\mathbf{F}$ .

**Theorem 4** *Suppose the conditions of Lemma 1 hold. Let  $m := \text{rank}(\widehat{\mathbf{F}})$ . Then*

$$i) (1 - \varepsilon) J^* \leq J_n^* \leq (1 + \varepsilon) J^*, \text{ and}$$

ii) if  $\mathbf{F}$  is linear, then with probability at least  $1 - m^{-\alpha}$

$$\|\mathbf{u}_n^* - \mathbf{u}^*\| \leq \frac{\varepsilon}{\sigma_{\min}^2(\mathbf{G})} \left( \|\widehat{\mathbf{F}}\| \|\mathbf{u}^*\| + \|\widehat{\mathbf{d}}\| \right) \|\widehat{\mathbf{F}}\|, \quad (34)$$

where  $\mathbf{G} := \left( \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{F}} + \mathbf{C}^{-1} \right)^{\frac{1}{2}}$ , and  $n = \mathcal{O}(\varepsilon^{-2} (2 + \alpha) \ln m)$ .

*Proof.* The first assertion follows from (25) and the definition of  $\mathbf{u}_n^*$  (10), indeed

$$J_n^* = J_n(\mathbf{u}_n^*) \leq J(\mathbf{u}^*) \leq (1 + \varepsilon) J(\mathbf{u}^*) = (1 + \varepsilon) J^*, \quad (35)$$

and the other direction is similar. For the second assertion, note that  $\mathbf{u}^*$  and  $\mathbf{u}_n^*$  are solutions of the following first optimality conditions

$$\left( \widehat{\mathbf{F}}^\top \widehat{\mathbf{F}} + \mathbf{C}^{-1} \right) \mathbf{u}^* = \widehat{\mathbf{F}}^\top \widehat{\mathbf{d}} + \mathbf{C}^{-1} \mathbf{u}_0, \quad (36a)$$

$$\left( \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{F}} + \mathbf{C}^{-1} \right) \mathbf{u}_n^* = \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{d}} + \mathbf{C}^{-1} \mathbf{u}_0. \quad (36b)$$

Define  $\Delta := \mathbf{u}^* - \mathbf{u}_n^*$ . An algebraic manipulation of (36) gives

$$\left( \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{F}} + \mathbf{C}^{-1} \right) \Delta = \left( \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{F}} - \widehat{\mathbf{F}}^\top \widehat{\mathbf{F}} \right) \mathbf{u}^* + \widehat{\mathbf{F}}^\top \widehat{\mathbf{d}} - \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{d}}. \quad (37)$$

Taking the inner product of both sides with  $\Delta$  we have

$$\begin{aligned} \left\langle \Delta, \left( \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{F}} + \mathbf{C}^{-1} \right) \Delta \right\rangle &= \left\langle \widehat{\mathbf{F}} \Delta, \Sigma \Sigma^\top \widehat{\mathbf{F}} \mathbf{u}^* - \widehat{\mathbf{F}} \mathbf{u}^* \right\rangle \\ &\quad + \left\langle \widehat{\mathbf{F}} \Delta, \widehat{\mathbf{d}} - \Sigma \Sigma^\top \widehat{\mathbf{d}} \right\rangle. \end{aligned} \quad (38)$$

Then we can bound the left-hand side of (38):

$$\left\langle \Delta, \left( \widehat{\mathbf{F}}^\top \Sigma \Sigma^\top \widehat{\mathbf{F}} + \mathbf{C}^{-1} \right) \Delta \right\rangle \geq \sigma_{\min}^2(\mathbf{G}) \|\Delta\|^2. \quad (39)$$

To bound terms on right hand side of (38), we need the following straightforward variant of (31), i.e.  $\forall \mathbf{v} \in \mathbb{R}^N$  and  $n = \mathcal{O}(\varepsilon^{-2})$ :

$$\|\Sigma \Sigma^\top \mathbf{v} - \mathbf{v}\| \leq \varepsilon \|\mathbf{v}\|. \quad (40)$$

Using the Cauchy-Schwarz inequality we have

$$\left\langle \widehat{\mathbf{F}} \Delta, \Sigma \Sigma^\top \widehat{\mathbf{F}} \mathbf{u}^* - \widehat{\mathbf{F}} \mathbf{u}^* \right\rangle \leq \varepsilon \|\widehat{\mathbf{F}}\|^2 \|\Delta\| \|\mathbf{u}^*\|, \quad (41a)$$

$$\left\langle \widehat{\mathbf{F}} \Delta, \widehat{\mathbf{d}} - \Sigma \Sigma^\top \widehat{\mathbf{d}} \right\rangle \leq \varepsilon \|\widehat{\mathbf{F}}\| \|\Delta\| \|\widehat{\mathbf{d}}\|, \quad (41b)$$

where we have used (40) and definition of matrix norm. Next, combining (41) and (39) ends the proof.

Note that for inequalities in (41) to be valid, it is sufficient to choose  $n, \alpha, \varepsilon$  such that (40) is valid for  $m$  basis vectors spanning the column space of  $\widehat{\mathbf{F}}$ , and hence  $n = \mathcal{O}(\varepsilon^{-2} (2 + \alpha) \ln m)$  by the union bound.  $\square$

**Remark 1** The bound in (34) is not a unique estimation. One can first rewrite  $J(\mathbf{u})$  and  $J_n(\mathbf{u}; \mathbf{r})$  as

$$\begin{aligned} J(\mathbf{u}) &= \frac{1}{2} \left\| \begin{bmatrix} \hat{\mathbf{d}} \\ \mathbf{C}^{-1/2} \mathbf{u}_0 \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{F}} \\ \mathbf{C}^{-1/2} \end{bmatrix} \mathbf{u} \right\|^2, \\ J_n(\mathbf{u}; \mathbf{r}) &= \frac{1}{2} \left\| \begin{bmatrix} \boldsymbol{\Sigma}^\top & 0 \\ 0 & \mathbf{I} \end{bmatrix} \left\{ \begin{bmatrix} \hat{\mathbf{d}} \\ \mathbf{C}^{-1/2} \mathbf{u}_0 \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{F}} \\ \mathbf{C}^{-1/2} \end{bmatrix} \mathbf{u} \right\} \right\|^2. \end{aligned}$$

If  $\boldsymbol{\Sigma}$  is a Johnson-Lindenstrauss transform, then  $\mathcal{S} := \begin{bmatrix} \boldsymbol{\Sigma}^\top & 0 \\ 0 & \mathbf{I} \end{bmatrix}$  is also a JLT with the same parameters:

$$\left\| \mathcal{S} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \right\|^2 = \|\boldsymbol{\Sigma} \mathbf{v}\|^2 + \|\mathbf{w}\|^2 \leq (1 + \varepsilon) \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 \leq (1 + \varepsilon) \left\| \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \right\|^2. \quad (42)$$

Applying [69, Theorem 12], we conclude that with probability at least  $1/3$ ,

$$\|\mathbf{u}_n^* - \mathbf{u}^*\| \leq \frac{\varepsilon}{\lambda_{\min}} \sqrt{J^*}, \quad (43)$$

where  $\lambda_{\min}$  is the minimum nonzero singular value of  $\begin{bmatrix} \hat{\mathbf{F}}^\top & \mathbf{C}^{-1/2} \end{bmatrix}^\top$ .

### 3.3. Data-scalability and cost complexity estimate

This section presents a qualitative discussion of the computational complexity and scalability of the randomized misfit approach. Numerical evidence of scalability to large data dimensions is presented in Section 4.3. For concreteness and ease of comparison, a Newton-type optimization method is assumed. The theory in Sections 3.1 and 3.2 is independent of the solver used.

The cost complexity of solving the randomized problem (9) is measured in number of PDE solves, i.e. solves of the forward or adjoint PDE and incremental variants. This characterization of complexity is agnostic to the specific governing forward PDE or PDE solver. For nontrivial forward problems, the total runtime of MAP point computation and uncertainty quantification is overwhelmingly dominated by the PDE solves; the cost of linear algebra is negligible in comparison [2, 14, 15, 21, 54].

In particular, with an inexact Newton-CG method, the cost of each Newton step is dominated by conjugate gradient (CG) iterations. Each CG iteration requires an application of the data misfit Hessian, which in turn requires a pair of incremental forward and adjoint PDE solves [2, 14, 15, 21, 70]. Thus the total work estimate is  $\mathcal{O}(2rk_{\text{Newton}})$  PDE solves. Here,  $k_{\text{Newton}}$  is the total number of Newton iterations and  $r$  is the numerical rank of the prior-preconditioned data misfit Hessian (or equivalently, the dimension of the likelihood-informed subspace (LIS) of parameter space [71]). Current

state-of-the-art implementations demonstrate that, for a wide class of inverse problems, the number of outer Newton iterations  $k_{\text{Newton}}$  and the numerical rank  $r$  are both independent of the mesh-size [15, 21, 70]. Mesh-independence is essential for ensuring scalability of a method to very high parameter dimensions.

The challenge is that even though  $r$  may be independent of the mesh, it still depends on the information content of the data. For many practical large-scale problems with high-dimensional data,  $r$  is on the order of hundreds or thousands (e.g.  $r = 5000$  in [72] and  $r = 1500$  for a linear 3D convection diffusion problem in [70]). Consequently, even with the best methods and modern supercomputers, solving the inverse problem is still computationally expensive.

Recall that for a given inverse problem,  $r$  is a fixed constant intrinsic to the misfit function, as it is the numerical rank of the prior-preconditioned Hessian of the misfit. A Newton-type method requires  $2r$  PDE solves (i.e.  $r$  inner iterations) at each outer iteration to sufficiently capture the  $r$  dominant modes of the misfit Hessian. Arbitrarily taking a much smaller number of inner iterations than  $r$  would result in more Newton iterations and degradation of the overall convergence. This constraint necessitates the use of a *surrogate misfit function*, with a Hessian that has numerical rank smaller than  $r$ , in order to bypass the impact of  $r$  on the overall cost of solving the inverse problem.

Ideally, this surrogate would leverage a small loss in the “level of parameter information in data” to obtain a large reduction in the overall computational cost of computing the inverse solution. In fact, this is what the randomized misfit approach can offer. The RMA cost is a surrogate cost that reduces the factor of  $r$  in the work estimate to an  $n \ll r$ , while providing a guarantee of solution viability. Note that the reduced misfit vector dimension  $n$  is a hard upper bound on the numerical rank of the misfit Hessian for the RMA cost  $J_n$  (9). This is numerically demonstrated for an elliptic inverse problem in Section 4.3.

Using the theory in Section 3.1, we can explicitly quantify the substantial gains in computational efficiency that are achieved with a specified accuracy level and a specified *confidence* level. This occurs by reframing the deterministic solution as one that holds with a given high probability.

The overall work estimate for the randomized misfit approach therefore is  $\mathcal{O}(2nk_{\text{Newton}})$ . This cost reduction analysis is markedly different from the analysis in the stochastic simultaneous source methods described in Section 1.1. By combining a large number of input sources  $s$  into a smaller number  $\tilde{s}$ , stochastic methods for multiple sources reduce the original problem from  $\mathcal{O}(2rk_{\text{Newton}}s)$  to  $\mathcal{O}(2rk_{\text{Newton}}\tilde{s})$  where  $\tilde{s} \ll s$ . Note that the RMA can provide a reduced work estimate in the most general class of inverse problems where  $s = 1$  and a guarantee of solution viability, whereas randomized simultaneous source methods cannot.



#### 4. Numerical experiments

In this section we demonstrate the randomized misfit approach with different distributions for  $\mathbf{r}$  in (9). We also verify that the convergence is indeed  $\mathcal{O}(1/\sqrt{n})$  as guaranteed by Theorem 3. Lastly we verify Theorem 2, the statistical Morozov's discrepancy principle.

The distributions that we test with the randomized misfit approach are:

- Gaussian
- Rademacher
- Achlioptas
- 95%-sparse ( $s$ -sparse (15) with  $s = 20$ )
- 99%-sparse ( $s$ -sparse (15) with  $s = 100$ )
- Uniform  $\mathcal{U} [-\sqrt{12}/2, \sqrt{12}/2]$

There are many other distributions suitable for RMA in the literature on Johnson-Lindenstrauss transforms that we do not consider, particularly the Subsampled Randomized Hadamard Transform of [73, 74] and its subsequent fast and sparse variants. These will be tested in future work.

We remark that subsampling (random subset) matrices are not proper random projection matrices and thus are not suitable for use in the RMA (9). Random source encoding methods often test subsampling matrices to reduce the dimension of the misfit [29, 29, 75]. For many inverse problems with identifiable structure in the data (e.g. 3-D hydraulic tomography [76]), subsampling can be extremely effective for reducing the computational burden of large observational datasets. However, in the RMA, subsampling down to misfit dimension  $n$  is equivalent to choosing  $\mathbf{r}_j$  from the canonical set  $\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$  *without* replacement. Therefore the set  $\{\mathbf{r}_j\}_{j=1}^n$  is not an i.i.d set. Similar to [27, 29, 75], our numerical results (omitted here) are poorer with subsampling matrices compared to results with proper random projections. This is consistent with the idea discussed in Section 1 that random projections are *geometry preserving transformations*, and can preserve the geometric relationship between any large observational data set and the parameter-to-observable map. Random subset matrices do not possess this property in general.

For our model problem we consider the estimation of a distributed coefficient in an elliptic partial differential equation. This Poisson-type problem arises in various inverse applications, such as the heat conductivity or groundwater problem, or in finding a membrane with a given spatially-varying stiffness.

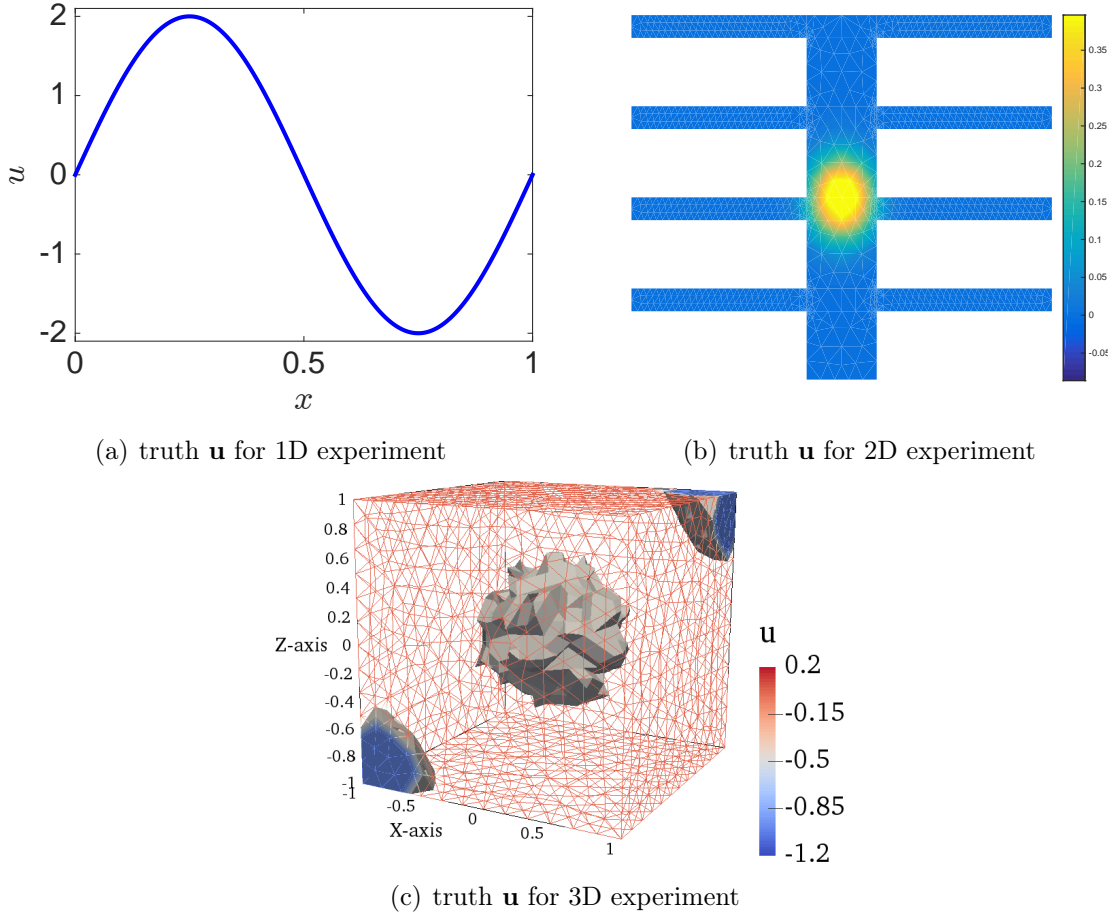
For concreteness we consider the heat conduction problem on an open bounded domain  $\Omega$ , governed by

$$\begin{aligned}
 -\nabla \cdot (e^{\mathbf{u}} \nabla w) &= 0 && \text{in } \Omega \\
 -e^{\mathbf{u}} \nabla w \cdot \mathbf{n} &= Bi w && \text{on } \partial\Omega \setminus \Gamma_R, \\
 -e^{\mathbf{u}} \nabla w \cdot \mathbf{n} &= -1 && \text{on } \Gamma_R,
 \end{aligned} \tag{44}$$

where  $\mathbf{u}$  is the logarithm of distributed thermal conductivity,  $w$  is the distributed forward state (temperature),  $\mathbf{n}$  is the unit outward normal on  $\partial\Omega$ , and  $Bi$  is the Biot number. Here,  $\Gamma_R$  is a portion of the boundary  $\partial\Omega$  on which the inflow heat flux is 1. The rest of the boundary is assumed to have Robin boundary condition. We are interested in reconstructing the distributed log conductivity  $\mathbf{u}$ , given noisy measurements of temperature  $w$  observed on  $\Omega$ .

The standard  $H^1(\Omega)$  finite element method is used to discretize the misfit and the regularization operator. The synthetic truths that we seek to recover are a 1-D sinusoidal curve, a 2-D Gaussian on a thermal fin, and a cube with nonzero log conductivity values on a sphere in the center and semispheres in the opposing corners. Figure 1 shows representations of  $\mathbf{u}_{\text{truth}}$  on a mesh for these cases.

The synthetic noisy temperature observations are then generated at all mesh points through the forward model (44). The misfit vector generated from 1(a) has data dimension  $N = 1025$  (with 1% percent added noise), from 1(b) has data dimension  $N = 1333$  (with .1% percent added noise), and from 1(c) has data dimension  $N = 2474$  (with .2% percent added noise), respectively.

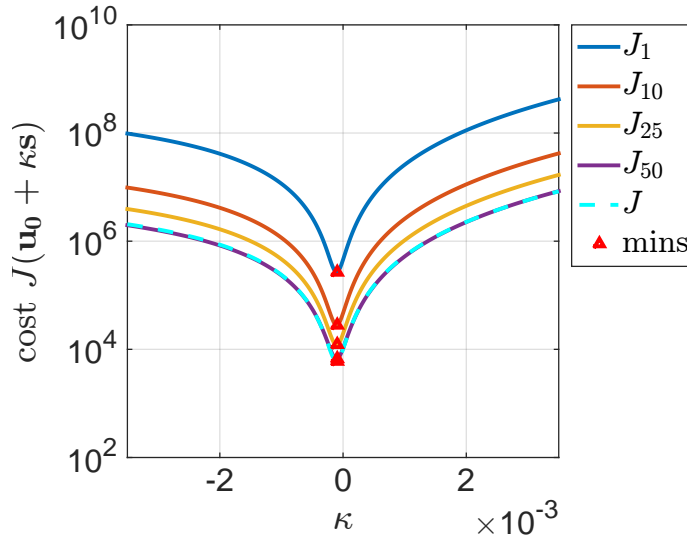


**Figure 1.** The distributed truth log conductivity parameters used in the experiments. The parameter fields are used to obtain noise-corrupted temperature data through the forward model (44).

For the inversion results we use an implementation of the trust region inexact Newton conjugate gradient method, for which some of the main ideas can be found in [60, 77–79]. Unless otherwise noted, the stopping criteria is when the Newton step size, cost function value, or norm of the gradient falls below  $10^{-6}$ .

#### 4.1. Convergence results

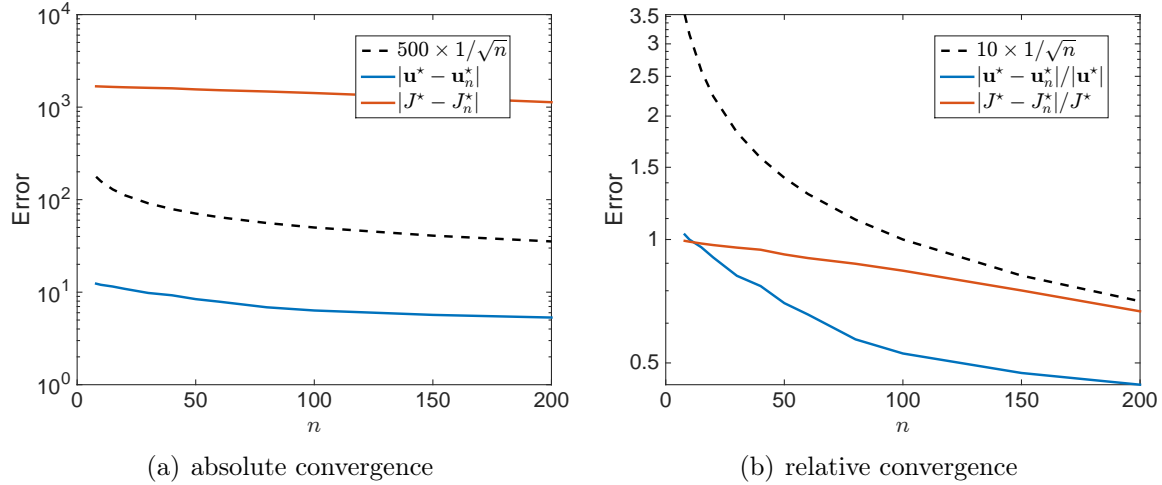
We first compare plots of the RMA cost  $J_n(\mathbf{u}_0)$  to the original cost  $J(\mathbf{u}_0)$  for a fixed distributed parameter  $\mathbf{u}_0$ , using the model heat problem (44). We choose a random  $\mathbf{u}_0$  from the prior distribution and construct the RMA cost  $J_n(\mathbf{u}_0)$  with the various random projections listed above. Since  $\mathbf{u}_0$  lives in high-dimensional space  $\mathbb{R}^m$ , where  $m$  is the number of finite element nodal values, for the purpose of visualization Figure 2 shows plots of the RMA cost  $\hat{J}_n(\kappa) := J_n(\mathbf{u}_0 + \kappa \mathbf{s})$  in a direction  $\mathbf{s} := \nabla J(\mathbf{u}_0)$  for the 3D example. For each of the random projections tested we observe convergence of  $\hat{J}_n(\kappa)$  to  $\hat{J}(\kappa)$  as  $n$  increases. More importantly, for all distributions, the minimizer of  $\hat{J}(\kappa)$  is well-approximated by  $\hat{J}_n(\kappa)$ , even for  $n$  small, as shown by Theorem 2. That is, although  $J_n$  for  $n = \mathcal{O}(1)$  is far from  $J$ , the local minimizers align. This is consistent with observed fidelity of randomized MAP points despite the slow convergence of the randomized cost, and similar phenomena seen in related methods. Plots with distributions other than Achlioptas and for the 1D and 2D examples are omitted when results are similar to the 3D Achlioptas experiments (see <http://users.ices.utexas.edu/~ellenle/RMAplots.pdf>).



**Figure 2.** Contours of the RMA cost  $J_n$  with different  $n$  versus the original cost  $J$ , for the 3D example with the Achlioptas distribution. Contours are evaluated along a 1D direction  $\mathbf{s}$  parameterized by  $\kappa$  and centered at a random parameter  $\mathbf{u}_0$  in the prior distribution. Red triangles indicate the minimum values of each contour.

Theorem 4 states that  $\mathbf{u}_n^*$ , the minimizer of  $J_n$ , and the minimum objective function value  $J_n^*$  converge at the same rate, given by the distortion tolerance  $\varepsilon$ , but with different

constants. Figure 2 illustrates how an RMA solution  $\hat{\mathbf{u}}_n^*$  may converge quickly to  $\hat{\mathbf{u}}^*$ , although convergence of the minimum value  $\hat{J}_n(\hat{\mathbf{u}}_n)$  to  $\hat{J}(\hat{\mathbf{u}})$  can be slow due to the different constant. To test this hypothesis at the actual minimizer  $\mathbf{u}^*$ , we plot the error of the RMA MAP point  $\mathbf{u}_n^*$  and its corresponding optimal value  $J_n^*$  in Figure 3 for the 3D example and the Achlioptas random projection $\ddagger$ . Data shown is the average of five runs. Both the absolute errors  $|J_n^* - J^*|$  and  $\|\mathbf{u}_n^* - \mathbf{u}^*\|$  and normalized errors  $|J_n^* - J^*|/|J^*|$  and  $\|\mathbf{u}_n^* - \mathbf{u}^*\|/\|\mathbf{u}^*\|$  are shown, and an  $\mathcal{O}(1/\sqrt{n})$  reference curve is plotted to show the convergence rate is indeed  $\mathcal{O}(1/\sqrt{n})$  for both  $\mathbf{u}_n^*$  and  $J_n^*$ . However, the absolute error of  $\mathbf{u}_n^*$  is orders of magnitude smaller than  $J_n^*$  for all considered random projections. Also, the relative error in  $\mathbf{u}_n^*$  decreases much faster than the relative error in  $J_n^*$  for  $1 < n < 100$ . Therefore a convergence analysis of the randomized cost  $J_n$  alone is not adequate for understanding the method efficacy in this range; the additional theory in Section 3.1 is required to characterize solution accuracy for  $n$  in the range of interest.

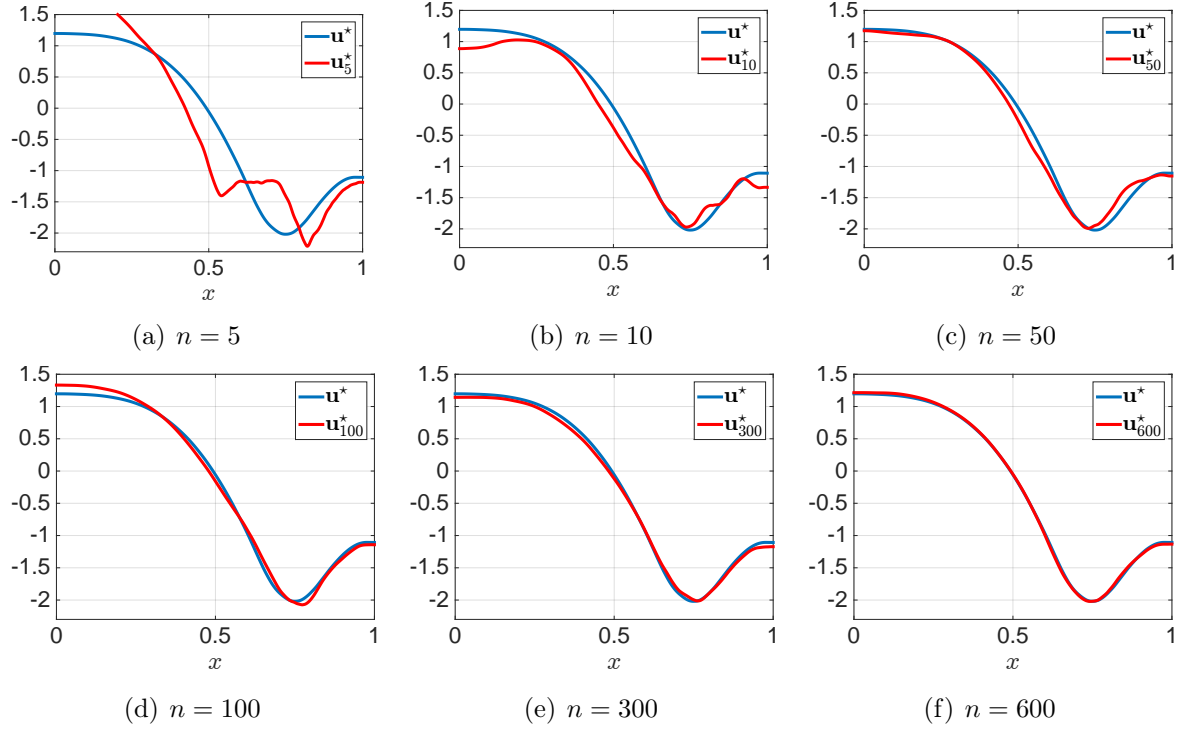


**Figure 3.** Log-linear plots of absolute errors  $|J_n^* - J^*|$  and  $\|\mathbf{u}_n^* - \mathbf{u}^*\|$ , relative errors  $|J_n^* - J^*|/|J^*|$  and  $\|\mathbf{u}_n^* - \mathbf{u}^*\|/\|\mathbf{u}^*\|$ , and  $\mathcal{O}(1/\sqrt{n})$  reference curves show the  $\mathcal{O}(1/\sqrt{n})$  convergence rate for both  $\mathbf{u}_n^*$  and  $J_n^*$  as given by Theorem 4.

Inversion results from minimizing the RMA cost with different  $n$  in the 1D, 2D, and 3D example are shown alongside the true MAP estimate  $\mathbf{u}^*$  in Figures 4, 5 and 6. The figures shown are results with  $\mathbf{r}$  distributed by the Achlioptas distribution (66% sparse). We see that the original MAP point  $\mathbf{u}^*$  is well-approximated by the RMA solution  $\mathbf{u}_n^*$  in all cases with  $50 \leq n \leq 100$ .

In a different experiment, we consider a 3D example in which only surface observations are available. The parameters are the same as the problem represented by Figure 1(c) but the data are now obtained from 901 observations on the surface of the cube (except the bottom surface), and the truth log conductivity is nonzero within the sphere of radius 0.5 centered at the origin as seen in Figure 7. Figure 8(d) shows

$\ddagger$  Again, similar results are seen with the 1D and 2D examples and with different random projections. They are omitted here.

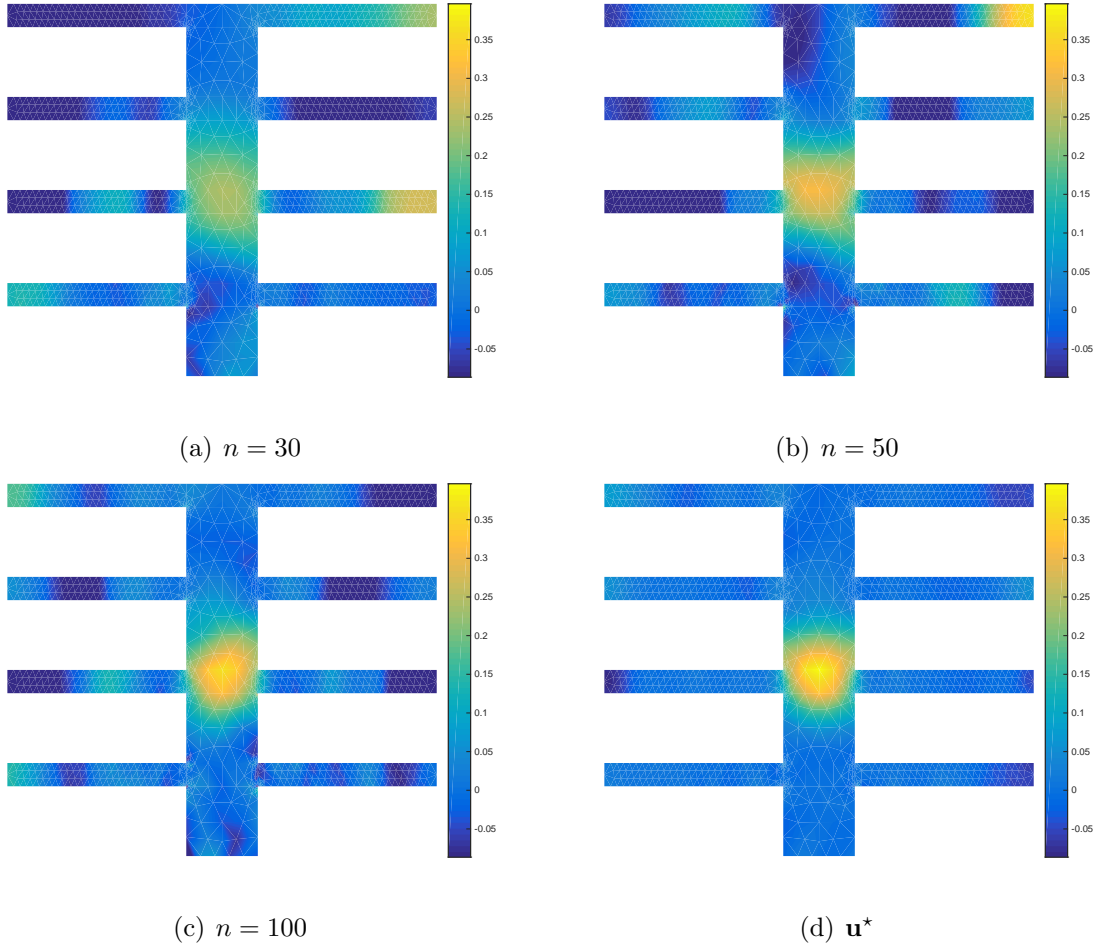


**Figure 4.** 1D elliptic PDE example: Convergence of  $\mathbf{u}_n^*$  to  $\mathbf{u}^*$  as  $n$  increases. The Achlioptas random projection (66% sparse) is used for  $\Sigma$  and the original data dimension is  $N = 1025$ .

the original MAP estimate  $\mathbf{u}^*$ . Compared to the above example the recovery is poorer, but this is expected due to having less observational data. Our interest however is in reducing the computational burden caused by the large data dimension while recovering a reasonable MAP estimation. Subsequently, we compare the RMA MAP point  $\mathbf{u}_n^*$  to the true MAP point  $\mathbf{u}^*$  (a minimizer of  $J$ ). The results in Figure 8 show the RMA solutions  $\mathbf{u}_n^*$  as  $n$  increases. As can be seen, with  $n = 150$ , i.e. a 6-fold reduction in the data misfit dimension, the RMA approximation  $\mathbf{u}_{150}^*$  is still a good approximation to the original MAP solution  $\mathbf{u}^*$ .

#### 4.2. Verification of Theorem 2

Table 1 presents results for solving the model problem for the 1D, 2D, and 3D examples with Morozov's criterion, again using the Achlioptas random projection in the randomized misfit approach. We perform several numerical experiments and choose an  $n$  for each example such that Morozov's principle is met for  $J_n(\mathbf{u}_n^*)$  with  $\tau' \approx 1$ . We then compute the corresponding ranges for  $\tau$  that are guaranteed with probability at least  $p \geq 1 - e^{-\beta}$ , after choosing an acceptable cost distortion tolerance of  $\varepsilon = 0.5$  and  $\beta$  as large as possible from (23). As can be seen, evaluating  $\mathcal{J}(\mathbf{u}_n^*)$  gives a  $\tau$  within the specified range, which satisfies Morozov's criterion. That is, even for moderately small values of  $n$ , if the discrepancy principle is satisfied for  $J_n(\mathbf{u}_n^*)$ , then the discrepancy principle is also satisfied for  $J(\mathbf{u}_n^*)$ . Thus  $\mathbf{u}_n^*$  is a discrepancy principle-satisfying solution



**Figure 5.** 2D elliptic PDE example: Convergence of  $\mathbf{u}_n^*$  to  $\mathbf{u}^*$  as  $n$  increases. The Achlioptas random projection (66% sparse) is used for  $\Sigma$  and the original data dimension is  $N = 1333$ .

for both the randomized reduced misfit dimension problem (9) and the original problem (6).

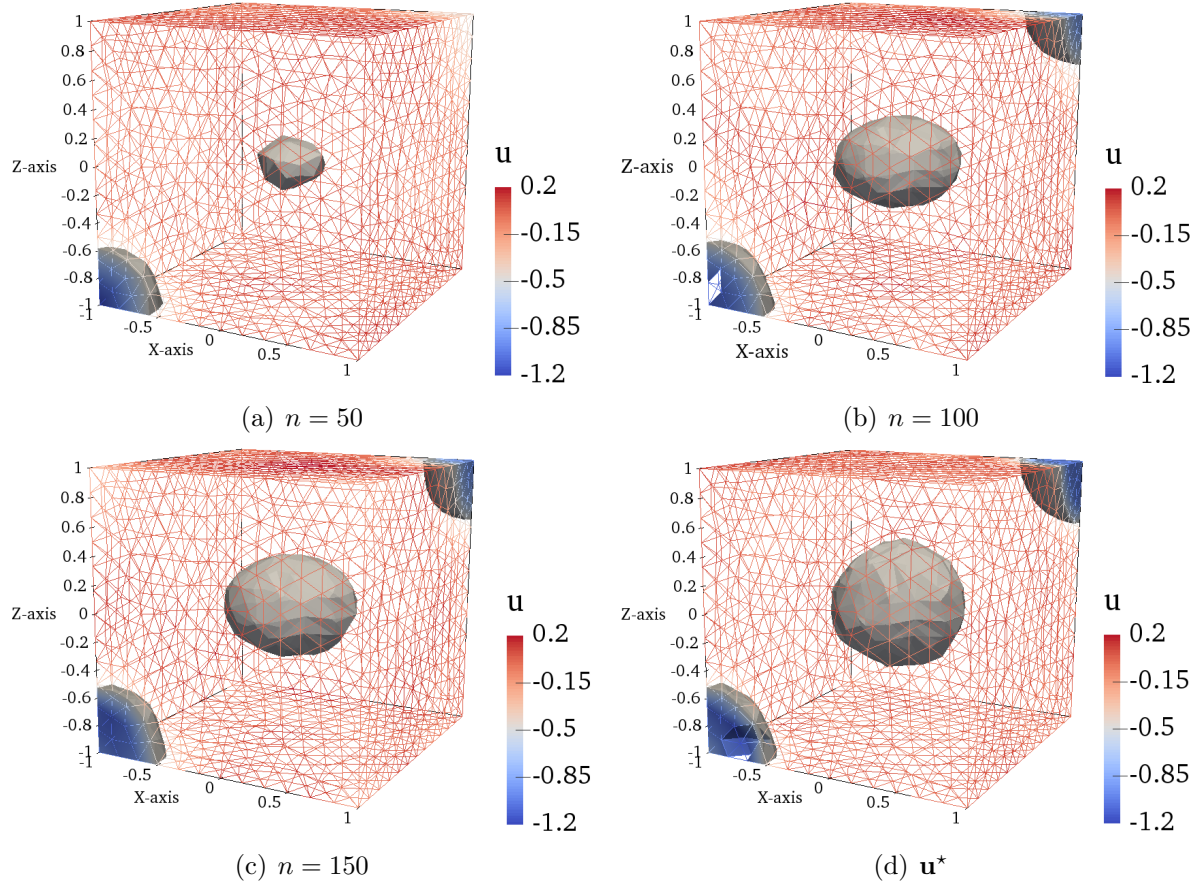
**Table 1.** Verification of Morozov's discrepancy principle for the RMA solution with  $\varepsilon = 0.5$ .

	$N$	$n$	$\mathcal{J}_n(\mathbf{u}_n^*)$	$\tau'$	$\left[\frac{\tau'}{1+\varepsilon}, \frac{\tau'}{1-\varepsilon}\right]$	$p$	$\mathcal{J}(\mathbf{u}_n^*)$	$\tau$
1D	1025	100	1220	1.190	$[0.793, 2.380]$	95.6%	1074	1.048
2D	1333	50	1240	0.930	$[0.620, 1.860]$	79.0%	1406	1.055
3D	2474	75	2646	1.070	$[0.713, 2.139]$	90.4%	3928	1.588

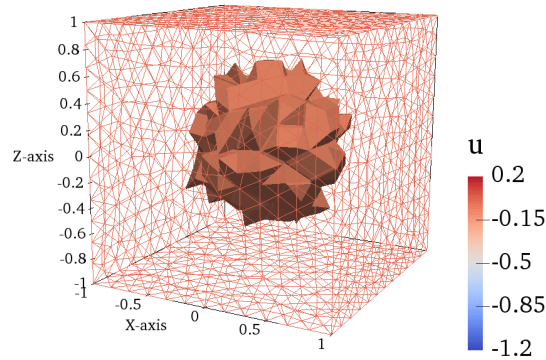
#### 4.3. Scalability and performance

We study the effect of the RMA reduced misfit dimension  $n$  on the overall algorithmic scalability of solving large-scale PDE-constrained inverse problems with high



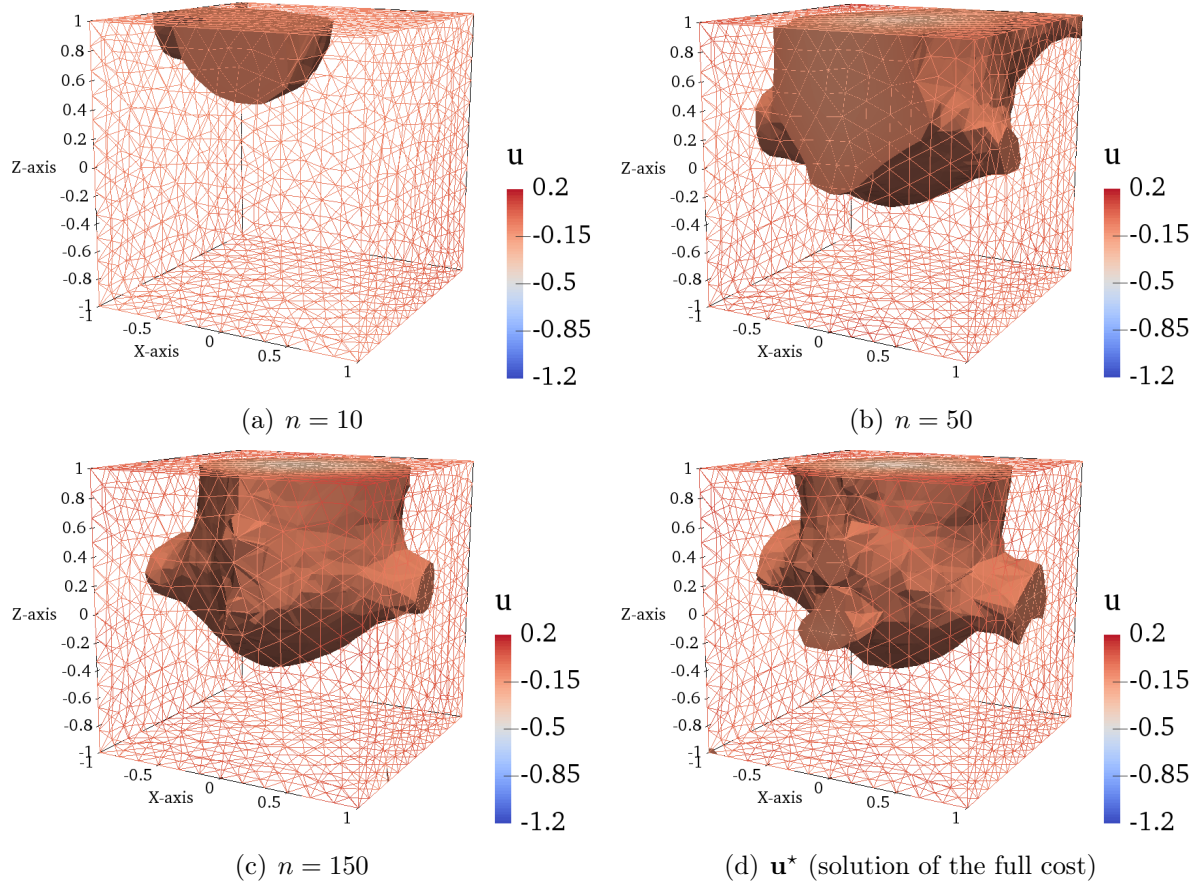


**Figure 6.** 3D elliptic PDE example: Convergence of  $\mathbf{u}_n^*$  to  $\mathbf{u}^*$  as  $n$  increases. The Achlioptas random projection (66% sparse) is used for  $\Sigma$  and the original data dimension is  $N = 2474$ .



**Figure 7.** Truth  $u$  for 3D experiment with surface observations: The same number of mesh elements as in Figure 1(c) is used but now the synthetic parameter is a single sphere, and observational data is obtained from  $N = 901$  mesh points on the top and side surfaces of the cube.

observational data dimensions. Specifically, we wish to show that RMA convergence is independent of  $r$ , the level of parameter information from the data (see Section 3.3). Figure 9 compares singular values of the prior-preconditioned misfit Hessian  $\mathbf{H}$  corresponding to the original problem cost  $J$  to the singular values of the surrogate

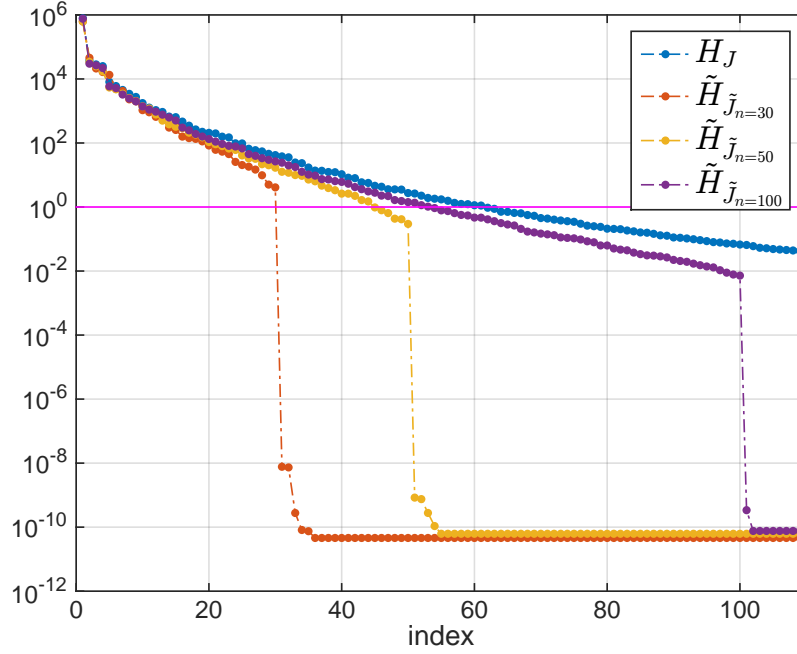


**Figure 8.** 3D elliptic PDE example with surface observations: Convergence of  $\mathbf{u}_n^*$  to  $\mathbf{u}^*$  as  $n$  increases. The MAP solution is nearly approximated with an RMA reduced misfit dimension of  $n = 150$  (a 6-fold reduction from the  $N = 901$  observational data points on the surface). The Achlioptas random projection (66% sparse) is used for  $\Sigma$ .

prior-preconditioned misfit Hessian  $\tilde{\mathbf{H}}$  corresponding to the surrogate RMA cost  $J_n$  for  $n = 30, 50$ , and  $100$ . The Hessians are each evaluated at the same random point chosen from the prior. Note that the RMA reduced misfit dimension  $n$  is a hard upper bound on the numerical rank of  $\tilde{\mathbf{H}}$ , where numerical rank is the number of singular values greater than some threshold  $\epsilon \leq 1$ . Note also the faster spectral decay of the singular values of  $\tilde{\mathbf{H}}$  compared to  $\mathbf{H}$ . Faster decay demonstrates that the action of  $\tilde{\mathbf{H}}$  on a vector can be captured with fewer modes than the action of  $\mathbf{H}$ , resulting in decreased overall work complexity as detailed in Section 3.3. Similar behavior is observed when the Hessians are evaluated at zero, at another random point, and at the full MAP point  $\mathbf{u}^*$ , thus the plots are omitted.

Tables 2 and 3 respectively present algorithmic performance of the original 2D and 3D elliptic problem compared to ten trials of the RMA with various distributions. To investigate the effect of choosing the randomized misfit dimension  $n < r \leq N$  on work complexity,  $n$  is chosen to be 50 for the 2D example where  $N = 1333$ , and  $n = 300$  for the 3D example where  $N = 2474$ . The Newton-CG solver is terminated when the gradient, cost, or step size falls below a tolerance of  $10^{-6}$ , or after 200 Newton iterations





**Figure 9.** Log-linear spectra of the prior-preconditioned misfit Hessian for the 2D elliptic example. Each spectrum is evaluated at the same random parameter  $\mathbf{u}$  drawn from the prior. Numerical rank is the number of singular values greater than some threshold  $\epsilon \leq 1$ . The misfit vector dimension ( $N$  in the original cost or  $n$  in the RMA cost) is a hard upper bound on the numerical rank of the prior-preconditioned misfit Hessian. The misfit dimension  $N$  for the original problem is 1333.

for the 2D example and 15 Newton iterations for the 3D example. Each trial uses a different random number generator seed. We observe that on average, using the RMA *with any distribution* results in close to half as many PDE solves compared to solving the full deterministic problem in the 2D example, and 14 to 28 percent fewer PDE solves in the 3D example. There appears to be little demonstrable difference in the quality of the reconstruction as well; all experiments are successful in reconstructing the Gaussian blob of high conductivity. Further investigation on very large problems ( $r = \mathcal{O}(1000)$  or larger) is needed.

**Table 2.** Comparison of cost complexity measured in total number of PDE solves needed to resolve the 2D elliptic problem with a Gauss-Newton solver. AVG is average over ten trials. Convergence tolerances for the cost, gradient, and step size are set to  $10^{-6}$  and the maximum number of Newton iterations allowed is 200.

#PDE SOLVES	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	AVG
Deterministic	2423	2423	2423	2423	2423	2423	2423	2423	2423	2423	<b>2423.0</b>
Rademacher	1303	1298	1273	1225	1279	1252	1270	1267	1252	1274	<b>1269.3</b>
Achlioptas	1149	1253	1293	1266	1253	1245	1267	1262	1231	1254	<b>1247.3</b>
95-percent sparse	1287	1272	1230	1273	1235	1217	1252	1248	1293	1238	<b>1254.5</b>
99-percent sparse	1212	1243	1245	1247	1250	1263	1268	1226	1218	1274	<b>1244.6</b>
Gaussian	1237	1258	1224	1226	1240	1273	1278	1255	1247	1234	<b>1247.2</b>
Uniform	1217	1244	1233	1242	1241	1264	1259	1262	1275	1248	<b>1248.5</b>

**Table 3.** Comparison of cost complexity measured in total number of PDE solves needed to resolve the 3D elliptic problem with a Gauss-Newton solver. AVG is average over ten trials. Convergence tolerances for the cost, gradient, and step size are set to  $10^{-6}$  and the maximum number of Newton iterations allowed is 15.

#PDE SOLVES	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	AVG
Deterministic	331	331	331	331	331	331	331	331	331	331	<b>331.0</b>
Rademacher	249	307	383	219	263	285	303	235	217	253	<b>271.4</b>
Achlioptas	259	287	263	257	283	203	259	223	319	255	<b>260.8</b>
95-percent sparse	279	273	221	231	285	241	255	313	217	367	<b>268.2</b>
99-percent sparse	277	321	271	249	279	395	273	223	281	279	<b>284.8</b>
Gaussian	219	223	323	199	257	233	229	257	211	225	<b>237.6</b>
Uniform	247	285	317	285	221	271	251	241	201	249	<b>256.8</b>

## 5. Conclusions and future work

A randomized misfit approach is presented for reducing computational complexity induced by big data in general large-scale PDE-constrained inverse problems. The method permits a novel analysis of the stochastic cost function and its minimizer via probabilistic bounds from random projection theory. It is shown that a subgaussian distribution guarantees the solution obtained from the randomized misfit approach will satisfy Morozov’s discrepancy principle with a low failure rate (that decays exponentially with respect to the reduced dimension  $n$ ).

It is shown that the stochastically derived method is equivalent to applying a random projection to the data misfit vector. This results in a stochastic programming-based proof (up to a constant) of a Johnson-Lindenstrauss lemma variant proved previously (see, e.g. [80, 81] for proofs based on combinatorics and communication theory, respectively). Our connection provides two main theoretical insights. The first is intuition into the surprising numerical accuracy with small reduced misfit dimension  $n$ . This phenomenon has been noted in related stochastic methods, particularly in random source encoding methods, without theoretical explanation. The second is an intuition into the ubiquitous  $\mathcal{O}(1/\sqrt{n})$  factor in Johnson-Lindenstrauss transforms (a rate shown to be tight by [80]) using a Monte Carlo framework.

The focus of this work is on the framework and resulting analysis of the method. We presented results for a medium size ( $N = \mathcal{O}(10^3)$ ) synthetic example in 1D, 2D, and 3D and different distributions for numerical justification of theoretical results and illustration of the method. Results presented here are valid for nonlinear inverse problems with the exception of part (ii) in Theorem 4 (which only applies to linear forward models). We expect such a result is also true for nonlinear inverse problems, and this is under investigation.

Combining dimension reduction and uncertainty quantification is the broader focus of our ongoing work towards developing scalable methods for large-scale inverse problems in high-dimensional parameter space with big data. Our current research includes an application of the randomized misfit approach to larger problems with big data, e.g. time-dependent data governed by expensive-to-solve forward models, and an extension

to the Bayesian solution. One project involves a large-scale multi-tracer test inverse problem governed by an expensive-to-solve reservoir simulation. Also in forthcoming tangential work we will compare different randomization frameworks for solving inverse problems.

## Acknowledgments

We thank the anonymous referees for their valuable comments, suggestions, and support. Their efforts helped improve the manuscript significantly. We would like to thank Prof. Mark Girolami for pointing out the similarity between randomized projections and the randomized misfit approach, which led to the connection with Johnson-Lindenstrauss theory. This in turn allowed us to carry out the analysis of the randomized misfit approach presented here. We also thank Vishwas Rao for careful proofreading. This research was partially supported by Department of Energy (DOE) grants DE-SC0010518 and DE-SC0011118. We are grateful for the support.

## References

- [1] Komatitsch D, Tsuboi S, Ji C and Tromp J 2003 A 14.6 billion degrees of freedom, 5 teraflops, 2.5 terabyte earthquake simulation on the Earth Simulator *SC03: Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis* (ACM/IEEE)
- [2] Bui-Thanh T, Burstedde C, Ghattas O, Martin J, Stadler G and Wilcox L C 2012 Extreme-scale UQ for Bayesian inverse problems governed by PDEs *SC12: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*
- [3] Martin J R 2015 *A computational framework for the solution of infinite-dimensional Bayesian statistical inverse problems with application to global seismic inversion* Ph.D. thesis The University of Texas at Austin
- [4] Aravkin A, Friedlander M P, Herrmann F J and Van Leeuwen T 2012 *Mathematical Programming* **134** 101–125
- [5] Moghaddam P P, Herrmann F J *et al.* 2010 Randomized full-waveform inversion: a dimensionality-reduction approach *2010 SEG Annual Meeting* (Society of Exploration Geophysicists)
- [6] van Leeuwen T, Aravkin A Y and Herrmann F J 2011 *International Journal of Geophysics* **2011**
- [7] Schraudolph N N and Graepel T 2003 Combining conjugate direction methods with stochastic approximation of gradients. *AISTATS*
- [8] Finkel J R, Kleeman A and Manning C D 2008 Efficient, feature-based, conditional random field parsing. *ACL* vol 46 pp 959–967
- [9] Nemirovski A, Juditsky A, Lan G and Shapiro A 2009 *SIAM Journal on Optimization* **19** 1574–1609
- [10] Shapiro A, Dentcheva D and Ruszczyński A 2009 *Lectures on Stochastic Programming: Modeling and Theory* (Society for Industrial and Applied Mathematics)
- [11] Kleywegt A J, Shapiro A and Homem-de Mello T 2002 *SIAM Journal on Optimization* **12** 479–502
- [12] Halko N, Martinsson P G and Tropp J A 2011 *SIAM Review* **53** 217–288
- [13] Martinsson P G, Rokhlin V and Tygert M 2011 *Applied and Computational Harmonic Analysis* **30** 47–68 ISSN 10635203
- [14] Alexanderian A, Petra N, Stadler G and Ghattas O 2016 *SIAM Journal on Scientific Computing* **38** A243–A272 (Preprint <http://dx.doi.org/10.1137/140992564>) URL <http://dx.doi.org/10.1137/140992564>
- [15] Isaac T, Petra N, Stadler G and Ghattas O 2015 *Journal of Computational Physics* **296** 348–368

- [16] Xiang H and Zou J 2015 *Inverse Problems* **31** 085008
- [17] Xiang H and Zou J 2013 *Inverse Problems* **29** 085008
- [18] Chaillat S and Biros G 2012 *Journal of Computational Physics* **231** 4403–4421
- [19] Lee J and Kitanidis P 2014 *Water Resources Research* **50** 5410–5427
- [20] Kitanidis P and Lee J 2014 *Water Resources Research* **50** 5428–5443
- [21] Bui-Thanh T, Ghattas O, Martin J and Stadler G 2013 *SIAM Journal on Scientific Computing* **35** A2494–A2523
- [22] Saibaba A K and Kitanidis P K 2015 *Advances in Water Resources* **82** 124–138 ISSN 03091708
- [23] Alexanderian A, Petra N, Stadler G and Ghattas O 2014 *SIAM Journal on Scientific Computing* **36** A2122–A2148
- [24] Bui-Thanh T and Girolami M A 2014 *Inverse Problems* **Special Issue** 114014
- [25] Saibaba A K, Lee J and Kitanidis P K 2015 *Numerical Linear Algebra with Applications* ISSN 1099-1506
- [26] Roosta-Khorasani F 2015 *Randomized algorithms for solving large scale nonlinear least squares problems* Ph.D. thesis University of British Columbia
- [27] Roosta-Khorasani F, Van Den Doel K and Ascher U 2014 *Electron. Trans. Numer. Anal* **42** 177–196
- [28] Roosta-Khorasani F and Ascher U 2015 *Foundations of Computational Mathematics* **15** 1187–1212
- [29] Roosta-Khorasani F, van den Doel K and Ascher U 2014 *SIAM Journal on Scientific Computing* **36** S3–S22
- [30] Roosta-Khorasani F, Székely G J and Ascher U M 2015 *SIAM/ASA Journal on Uncertainty Quantification* **3** 61–90
- [31] Routh P S, Lee S, Neelamani R, Krebs J R, Lazaratos S and Marcinkovich C 2014 Simultaneous source encoding and source separation as a practical solution for full wavefield inversion uS Patent 8,775,143
- [32] Krebs J R, Anderson J E, Neelamani R, Jing C, Hinkley D, Dickens T A, Krohn C E and Traynin P 2012 Iterative inversion of data from simultaneous geophysical sources uS Patent 8,121,823
- [33] Neelamani R, Krohn C E, Krebs J R, Romberg J K, Deffenbaugh M and Anderson J E 2010 *Geophysics* **75** WB15–WB27
- [34] Krebs J R, Anderson J E, Hinkley D, Neelamani R, Lee S, Baumstein A and Lacasse M D 2009 *Geophysics* **74** WCC177–WCC188
- [35] Haber E, Chung M and Herrmann F 2012 *SIAM Journal on Optimization* **22** 739–757
- [36] Haber E, Ascher U M and Oldenburg D W 2004 *Geophysics* **69** 1216–1228
- [37] Oldenburg D W, Haber E and Shekhtman R 2012 *Geophysics* **78** E47–E57
- [38] Pratt R G 1999 *Geophysics* **64** 888–901
- [39] Virieux J and Operto S 2009 *Geophysics* **74** WCC1–WCC26 URL <http://link.aip.org/link/?GPY/74/WCC1/1>
- [40] Herrmann F J, Erlangga Y A and Lin T T 2009 *Geophysics* **74** A35–A40
- [41] Haber E, Heldmann S and Ascher U 2007 *Inverse Problems* **23** 1659–1676
- [42] Duraiswami R, Sarkar K and Chahine G L 1998 *Engineering Analysis with Boundary Elements* **22** 13–31
- [43] Kaipio J and Somersalo E 2005 *Statistical and Computational Inverse Problems (Applied Mathematical Sciences vol 160)* (New York: Springer-Verlag)
- [44] Young J and Ridzal D 2012 *SIAM Journal on Scientific Computing* **34** A2344–A2365
- [45] Hutchinson M F 1990 *Communications in Statistics-Simulation and Computation* **19** 433–450
- [46] Achlioptas D 2003 *Journal of Computer and System Sciences* **66** 671–687 ISSN 00220000
- [47] Dirksen S 2015 *Foundations of Computational Mathematics*
- [48] Matousek J 2008 *Random Struct. Algorithms* **33** 142–156
- [49] Indyk P and Naor A 2007 *ACM Transactions on Algorithms (TALG)* **3** 31
- [50] Bui-Thanh T and Ghattas O 2012 *Inverse Problems* **28** 055001
- [51] Bui-Thanh T and Ghattas O 2012 *Inverse Problems* **28** 055002

- [52] Bui-Thanh T and Ghattas O 2013 *Inverse Problems and Imaging* **7** 1139–1155
- [53] Stuart A M 2010 *Acta Numerica* **19** 451–559
- [54] Petra N, Martin J, Stadler G and Ghattas O 2014 *SIAM Journal on Scientific Computing*
- [55] Dashti M, Harris S and Stuart A 2012 *Inverse Problems and Imaging* **6** 183–200
- [56] Lassas M, Saksman E and Siltanen S 2009 *Inverse Problems and Imaging* **3** 87–122
- [57] Bui-Thanh T and Ghattas O 2015 *Inverse Problems and Imaging* **9** 27–53
- [58] Lindgren F, Rue H and Lindström J 2011 *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73** 423–498 ISSN 1467-9868 URL <http://dx.doi.org/10.1111/j.1467-9868.2011.00777.x>
- [59] Kitanidis P K 2012 *Advances in Water Resources* **36** 3–10
- [60] Branch M A, Coleman T F and Li Y 1999 *SIAM Journal on Scientific Computing* **21** 1–23 (electronic)
- [61] Touchette H 2009 *Physics Reports* **478** 1–69
- [62] Kelly F P 1991 *Queueing systems* **9** 5–15
- [63] Li P, Hastie T J and Church K W 2006 *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '06* 287
- [64] Stroock D W 2011 *Probability Theory: An Analytic View* 2nd ed (Cambridge University Press, Cambridge)
- [65] Holub V and Fridrich J 2013 *Information Forensics and Security, IEEE Transactions on* **8** 1996–2006
- [66] Liu L, Fieguth P, Clausi D and Kuang G 2012 *Pattern Recognition* **45** 2405–2418
- [67] Fowler J E and Du Q 2012 *Image Processing, IEEE Transactions on* **21** 184–195
- [68] Morozov V A 1966 *Soviet Math. Dokl.* **7**
- [69] Sarlos T 2006 *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)* 143–152 ISSN 0272-5428
- [70] Flath H P, Wilcox L C, Akçelik V, Hill J, van Bloemen Waanders B and Ghattas O 2011 *SIAM Journal on Scientific Computing* **33** 407–432
- [71] Cui T, Martin J, Marzouk Y M, Solonen A and Spantini A 2014 *Inverse Problems* **30** 114015
- [72] Isaac T, Petra N, Stadler G and Ghattas O 2015 *Journal of Computational Physics* **296** 348–368
- [73] Ailon N and Chazelle B 2009 *SIAM Journal of Computing* **39** 302–322
- [74] Tropp J A 2010 *Advances in Adaptive Data Analysis* **03** 8 ISSN 1793-5369
- [75] van den Doel K and Ascher U M 2012 *SIAM Journal on Scientific Computing* **34** A185–A205
- [76] Cardiff M, Barrash W and Kitanidis P K 2013 *Water Resources Research* **49** 7311–7326
- [77] Coleman T F and Li Y 1996 *SIAM Journal on Optimization* **6** 418–445
- [78] Bui-Thanh T 2007 *Model-Constrained Optimization Methods for Reduction of Parameterized Large-Scale Systems* Ph.D. thesis Department of Aeronautics and Astronautics, MIT
- [79] Nocedal J and Wright S J 2006 *Numerical Optimization* 2nd ed (Berlin, Heidelberg, New York: Springer Verlag)
- [80] Alon N 2003 *Discrete Mathematics* **273** 31–53
- [81] Jayram T and Woodruff D P 2013 *ACM Transactions on Algorithms (TALG)* **9** 26